

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 16:15:20 ; Search time 4236 Seconds

(without alignments)  
10426.099 Million cell updates/sec

Title: US-10-807-228-2

Perfect score: 1212  
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACATCATCGCAAC 1212

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing filter 45 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_hnc:\*  
4: gb\_esc3:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	322.6	26.6	470	4	B1376242 BFLG3_000
2	221	18.2	470	1	A1906433 IL-BT109-
3	207.8	17.1	494	6	CD307119 StrPu691.
4	188.6	15.6	442	6	CD295681 StrPu691.
5	123.6	10.2	519	6	CD296311 StrPu691.
6	83.2	6.9	551	7	CF919042 BFL0531.
7	83.2	6.9	553	4	B1387857 BFL26_002
8	69.6	5.7	524	8	B2894814 Hg4_0102
9	69.2	5.7	722	6	CB656525 OSJNEC10P
10	67.6	5.6	722	6	CB683938 OSJNEC10P
11	63.6	5.2	925	9	CNS0091P AL053013 Drosophila
12	60.8	5.0	712	6	CD869465 AZ02_111L
13	60.8	5.0	812	6	CB659172 OSJNEC10
14	60.6	5.0	604	1	AU162766 AU162766
15	60.6	5.0	668	1	CD225047 CCC1_37_H
16	60.4	5.0	512	4	BM140351 WHE0474_a
17	60.2	5.0	719	4	CF487085 POL1_41_D
18	60	5.0	748	7	CN143339 WOUND1_15
19	60	5.0	768	7	CN148454 WOUND1_56
20	60	5.0	814	7	CN126847 RH0H1_19
21	60	5.0	1115	7	CK208301 FGAS02000
22	59.8	4.9	635	6	CA254345 SCBPFL411
23	59.8	4.9	707	7	CF874011 tric035xc
24	59.8	4.9	732	7	CF866540 tric030xc

25	59.8	4.9	748	6	CA227320	CA227320	SCULF301
26	59.8	4.9	762	6	CB903555	CB903555	tric035xc
27	59.8	4.9	802	6	CB902534	CB902534	tric030xc
28	59.6	4.9	683	6	CA269770	CA269770	SCMCR308
29	59.4	4.9	670	6	CD429468	ETH1_4_D0	CD429468
30	59	4.9	579	2	BE040798	OF1G11_O	BE040798
31	59	4.9	690	2	BE041110	OP19H01_O	BE041110
32	59	4.9	806	7	CN127041	RH0H1_20	CN127041
33	58.8	4.9	888	9	CG440030	OCVPH10TH	CG440030
34	58.6	4.8	532	4	BG241153	OV1_38_H0	BG241153
35	58.6	4.8	635	6	CD223689	CCCL_29_A	CD223689
36	58.6	4.8	641	7	CF481017	POL1_69_B	CF481017
37	58.6	4.8	646	6	CD229216	CCCL_13_A	CD229216
38	58.6	4.8	660	6	CD227467	CCCL_51_F05	CD227467
39	58.6	4.8	664	7	CN133040	OX1_5_F05	CN133040
40	58.6	4.8	666	6	CD222873	CCCL_24_F	CD222873
41	58.6	4.8	667	7	CF485910	POL1_34_B	CF485910
42	58.6	4.8	672	6	CD224549	CCCL_34_E	CD224549
43	58.6	4.8	680	7	CF432900	NIT1_19_E	CF432900
44	58.6	4.8	703	7	CF487230	POL1_42_C	CF487230
45	58.6	4.8	710	7	CF488266	POL1_48_F	CF488266

#### ALIGNMENTS

RESULT 1  
B1376242  
Locus  
DEFINITION  
BFLG3\_000038 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG3 or MPMGP498) Branchiostoma floridae cDNA clone MPMGP498A1518 5', mRNA sequence.

ACCESSION  
B1376242  
VERSION  
B1376242.1 GI:30911206  
KEYWORDS  
EST.  
SOURCE  
Branchiostoma floridae  
ORGANISM  
Branchiostoma floridae  
Bukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

REFERENCE  
1 (bases 1 to 623)  
Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.  
New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL  
Genome Res. 13 (6A), 1056-1066 (2003)  
MEDLINE  
22683279  
PUBMED  
12799346  
COMMENT  
Contact: Panopoulou G  
Laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhestr.63-73, D-14195 Berlin, Germany  
Tel.: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoulou@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One cluster per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)  
PCR primers  
FORWARD: 5' CCCAGGCTTACCTTATGCTCCGCTCG 3' (M13SP)  
BACKWARD: 5' GCTATTACGACGACGCGAAGGGATGTG 3' (M13SP)  
Insert length: 1200 Std Error: 0.00  
Seq primer: 5'-CCGCGCGAATCCCGGCT-3', pSPORT3/86  
High quality sequence stop: 623.

## FEATURES

Location/Qualifiers

1. 623

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMP498A1518"

/issue\_type="whole embryo"

/dev\_stage="5-6 hrs (gastrula stage)"

/lab\_host="E.coli, XL1 blue"

/clone\_lib="Amplioxus 5-6 hrs cDNA library (Name convention: BFL6 or MPMP498)"

/note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); Oligodt primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTCTGATCGGAGCGCGCCG (T)15-3' and a SalI 5'-TCGACCCGCGCTCCG-3' adapters (Gibco BRL)."

## ORIGIN

Query Match 26.6%; Score 322.6; DB 4; Length 623;  
Best Local Similarity 70.4%; Pred. No. 1.4e-60;  
Matches 430; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 570 GCCGAGCATGAGTGGCGATCCGACCAATGCGATGATCCGAGATCCCAATC 629  
DB 13 GCGTAGTACGAGCTGGCGATGACCGCTCCCAACATGATACGGGAATGCCGAGC 72  
QY 630 GTTCCCTTCTGAGAGCTGATGACACCTGAGCTTCTGAGTCCGAGCATCAACCA 689  
DB 73 CTAACCGCATGGGGAATCATGACACTTGGGCTGATCCAGTCAAGGATCAACACTGA 132  
QY 690 CGGCGGCAATCCGCGTACCAACCGCATCGGCAATCCGCGATCTTTCGCTCA 749  
DB 133 CGGAGGACCAACCTCTGACGTCACGCGCGCGGCGAGATCTGTCATCA 192  
QY 750 CACCTTCCGAGATCTTGGCTACTACACCGGCTGAGAGCGACCGCTTCTGCGACCA 809  
DB 199 CACCTTCCGAGATCTTGGCTACTACACCGGCTGAGAGCGACCGCTTCTGCGACCA 252  
QY 810 TGTCCATGACCGAGCTTCTGATCTGAGAGAGAACTGGCCGCTGATCCGCGGCT 869  
DB 253 CGCTTCCGAGACCACTCTCGACTTGGGAGATCACTTAAGTATACCGCGCGGCT 312  
QY 870 CGAGCTGATCAAGCGCGCGCGCGCTGCAAGAGCATCGCATCGAGCTCAAGAGATGA 929  
DB 313 GAGAGTATCCGCGCGGAGCGCGCTGCTGTCATCGCACTGACCTGAACGAGATGA 372  
QY 930 CGCGAGTGGAGACTGCTGTAAGTACCGCTCTTGGCTATGCGCACTCTTGGCGTCT 989  
DB 373 CGGAGGACCAACCGGCTGCTGATACCGCGCATCGGCTACCTTGGCGTCT 432  
QY 990 GTTCCATCTACGAGTCTGCGAGGCGGCGTGAAGTCTGCGAGAGACATGACACCGAGCT 1049  
DB 433 GTTCCATCTACGAGTCTGCGAGGCGGCGTGAAGTCTGCGAGAGAGTCTGATACCGTCT 492  
QY 1050 GAAAGCGCGCATGAGTCTGATGAGAGCGGATGAGTCTGCGAGAGAGAGAGAGAGAGAG 1109  
DB 493 GAAAGCGCGCATGAGTCTGATGAGAGCGGATGAGTCTGCGAGAGAGAGAGAGAGAGAG 552  
QY 1110 TCCCGCGCGCTATCGGAGACGACATCTGATCTGCGGAGAGAGAGAGAGAGAGAGAG 1169  
DB 553 ACCGCGCGCTTACCGGAGACGACATCTTGTGTGATGATGAGAGAGAGAGAGAGAGAG 612  
QY 1170 CACCGGCTTCC 1180  
DB 613 TACAGGCTCC 623

RESULT 2  
AI906433/c AI906433 470 bp mRNA linear EST 30-MAR-2000  
LOCUS IL-BT109-280199-002 BT109 Homo sapiens cDNA, mRNA sequence.  
DEFINITION AI906433  
ACCESSION AI906433  
VERSION AI906433.1 GI:6496820

## KEYWORDS

EST.  
Homo sapiens (human)Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.REFERENCE  
AUTHORS  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brito, M. R.,  
Nagai, M. A., da Silva, W. J., Zago, M. A., Bordini, S., Costa, P. F.,  
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,  
Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V.,  
O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and  
Simpson, A. J.TITLE  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Rui Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?it=il-BT109-002.html  
kt3=280199&ct=1)  
Seq primer: puc 18 forward.

## FEATURES

Location/Qualifiers

1. 470  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stage="Adult"  
/clone\_lib="BT109"/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

## ORIGIN

Query Match 18.2%; Score 221; DB 1; Length 470;  
Best Local Similarity 69.4%; Pred. No. 3.2e-38;  
Matches 313; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

QY 630 GTTCCCTTCTGAGAGCTGATGACACCTGACCTGTTTCAAGTGGGATCAACACCA 689  
DB 455 GATCCCTGATGGGCGTGAATACACCTGATGACCTGTTTCAAGTGGGATCAACACCA 396  
QY 690 CGGCGGCAATCCGCTGACCAACCGCATGTCATCCGCGGACATCTTTCCTCA 749  
DB 395 TGGGCGCATACCGCGGTGACAGCGCGCGGTGGGAGAGAGAGAGAGAGAGAGAG 795  
QY 750 CACCTTCCGAGTATCTTTCGCGG- TACTACACCGCGGTGGAGAGAGAGAGAGAGAGAG 808  
DB 335 TTGCTTCCGAGTATCCGCGGATATTAACCGCGGTGGAGAGAGAGAGAGAGAGAG 868  
QY 809 ATGTGATGACCGACGCTGACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868  
DB 275 ACTGCGCGAGAGAGTACGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216  
QY 869 TCGACTGATCAAGCGCGCGCGCGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928  
DB 215 TGAAGTGTGCGCGCGCGAGTGGCTGACAGATGCTGCGGAGATGAGAGAGAGAG 156  
QY 929 ACCGAGATGGAGACTGCTGAAGTACCGCTCTTGGCTATGGGCACTCTTGGCGGTGC 988

Db 155 TCTGTGGCCAGCAGCTGTGTGACGACCGACTTGGGTATGAGCACTGCTTGGAGCT 96

Qy 988 TGTGCCACTACTAAGGTGCGGAGCGCGGTGAGTGGCGAGGACATCGACACCGAGC 1048

Db 95 TTAGCCACTACTAAGGTGCGGAGCGCGGTGAGTGGCGAGGACATCGACACCGTGC 36

Qy 1049 TGAAGCCCGGCAATGGTGTCTCCATGAGCC 1079

Db 35 TCGAACCGGGCAATGGTGTCTCATCGCCGC 5

RESULT 3  
CD307119/c 494 bp mRNA linear EST 16-SRP-2003

LOCUS Strp691.009486 Sea urchin larva cDNA library MPM691

DEFINITION Strongylocentrotus purpuratus cDNA clone

ACCESSION MPM691E0990/MP1\_SURUDI\_90E9 5', mRNA sequence.

VERSION CD307119

KEYWORDS CD307119.1 GI:34752168

SOURCE Strongylocentrotus purpuratus

ORGANISM Strongylocentrotus purpuratus

REFERENCE Pouskka, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A., Reinhardt, R., Herwig, R., Panopoulou, G., and Lehrach, H. Generation, annotation, evolutionary analysis, and database integration of 20,000 unique sea urchin EST clusters Genome Res. 13 (12), 2736-2746 (2003)

COMMENT JOURNAL Contact: Pouskka AJ Laboratory 145, dept. Lehrach Max-Planck-Institut fuer Molekulare Genetik Ihmestr.-63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: pouskka@molgen.mpg.de

REFERENCE The library was characterised by oligonucleotide fingerprinting (ONF) to reduce sequencing redundancy. According to the ONF procedure, clones that display the same hybridisation matrix with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per ONF cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of the other clones assigned to the same ONF cluster as the clone from which the above EST is generated is available at the sea urchin project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA clones and filters are distributed via the Resource Center/Primary Database of the German Human Genome Project (<http://www.tzpd.de>)

PCR Primers FORWARD: 5' CCCAGGCTTACACTTATGCTTCGGCTCG 3' (M13RSP) 5'-seq BACKWARD: 5' GCTATTACCGCAGCTGGGAGAGGAGGATG 3' (M13SP) 3'-seq Seq primer: 5'-CCGGTCCGGAATCCCGGCT-3' pSPORT3/86 High quality sequence stop: 494.

FEATURES  
Location/Qualifiers  
1..494  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="MPM691E0990/MP1\_SURUDI\_90E9"  
/cissue\_type="whole larva"  
/cissue="larva 2-3 weeks"  
/dev\_stage="B.coli, XLI blue"  
/clone\_lib="Sea urchin larva cDNA library MPM691"  
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; Random primed and directionally cloned in pSPORT15 using a NotI (5'-pGACTGCTTCTAGATCGGACGCGCCGCTT)15-3' and a SalI 5'-TCGACCCACGCGCTCG-3' adapters (Gibco BRL)"

ORIGIN  
Query Match 17.1%; Score 207.8; DB 6; Length 494;  
Best Local Similarity 66.8%; Pred. No. 2.6e-35;

Matches 312; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

Qy 730 GCGCAGATCTTCTGCTCAACACCTTCCGATGATCTTGGTACTACACCGCGCTGGAG 789

Db 476 GGTGACGGCATCTGCTTTACGCTTCCCTTATGATAGCGCGCTATTTCTCGGCGCTTGA 417

Qy 790 GCGACGCTGTTCTGCGACATGTCATGACCGCAC---CTGACATCTGGGAGGAAGAC 846

Db 416 CGAACCTTTTCTTAAACGACGTGCTTCCAGCCGCATCTTGAAGTGTGGACGCTCAAC 357

Qy 847 GTGCGCTGATTCGCGCGGCTCGAGCTGATCAAGCCGCGCGCTGCAAGGACATC 906

Db 356 TCGACGCTTCAACCGCGCGGATGAGGCTCATCAAGCGCTGATCAAAATGCTGCCAGCTC 297

Qy 907 GCCATCGAGCTCAACGAGATGATCCGCAATGGGACCTGCTGAATGACCGCTCTTGGC 966

Db 296 GCGCTCGAGTTGAACGAGATGATACCGGAGGAACCTTACATCAATCAAGACCTTCGCGC 237

Qy 967 TATGCGCACTGCTTCCGCGGTGCTGCGACTACTACGAGTCGCGGAGCGCGCTGAGCTG 1026

Db 236 TACGGGCACTCTTGGGGTCTCTATGCACTACGAGCGCGGACACTTTGAGCTTG 177

Qy 1027 CGCGAGACATCAACACCGAGCTGAAGCCCGCATGTGTCTCCATGAGCCGATGTG 1086

Db 176 AGGGAAGACATGAGACCTGTATACAGCTGTGTATGCTTCCATGAGGCACATCTC 117

Qy 1087 ATGCTGCGGAGGAGCATCCCGGTGCGCGGCTATCCGAGACAGACATCTGATCGTC 1146

Db 116 ACGATACGAGACGCGACACTGTGTGCAAGAGGCTACAGGAGCATGATATCATGTGTG 57

Qy 1147 GGGAGGAGCGGTCCGAGACATCACCGGCTTCCGCTTCGCGCGGA 1193

Db 56 ACTGAAGGGGCGCTGCTTCAATACCGGATTTCTTACGACCGGA 10

RESULT 4  
CD295681 442 bp mRNA linear EST 16-SRP-2003

LOCUS Strp691.007076 Sea urchin larva cDNA library MPM691

DEFINITION Strongylocentrotus purpuratus cDNA clone

ACCESSION MPM691E0990/MP1\_SURUDI\_90E9 3', mRNA sequence.

VERSION CD295681

KEYWORDS CD295681.1 GI:34746758

SOURCE Strongylocentrotus purpuratus

ORGANISM Strongylocentrotus purpuratus

REFERENCE Pouskka, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A., Reinhardt, R., Herwig, R., Panopoulou, G., and Lehrach, H. Generation, annotation, evolutionary analysis, and database integration of 20,000 unique sea urchin EST clusters Genome Res. 13 (12), 2736-2746 (2003)

COMMENT JOURNAL Contact: Pouskka AJ Laboratory 145, dept. Lehrach Max-Planck-Institut fuer Molekulare Genetik Ihmestr.-63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: pouskka@molgen.mpg.de

REFERENCE The library was characterised by oligonucleotide fingerprinting (ONF) to reduce sequencing redundancy. According to the ONF procedure, clones that display the same hybridisation matrix with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per ONF cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of the other clones assigned to the same ONF cluster as the clone from which the above EST is generated is available at the sea urchin project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA clones and filters are distributed via the Resource Center/Primary

Database of the German Human Genome Project (<http://www.rzpd.de>)

PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGCGAGCGGAGGAGTG 3' (M13FSP) 3'-seq  
Seq primer: 5' GCTATTACGCGAGCGGAGGAGTG 3' (M13FSP)  
High quality sequence stop: 442.  
Location/Qualifiers

## FEATURES

source

1..442  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="MPMG691E0990;MPI\_SURUDI\_90E9"  
/cissue\_type="whole larva"  
/dev\_stage="larva 2-3 weeks"  
/lab\_host="E.coli, XLI blue"  
/clone\_lib="Sea urchin larva cDNA library MPMG691"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-TGACGATGTTAGTGGAGCGGCGCC (T)15-3' and a  
SalI 5'-TCGACCCAGCGCTCCG-3' adapters (Gibco BRL)"

## ORIGIN

Query Match 15.6%; Score 188.6; DB 6; Length 442;  
Best Local Similarity 67.3%; Pred. No. 4.4e-11;  
Matches 282; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 730 GCGGACATCTTTCGCTCAACACCTTCGGATATCTTGGCTACTACACCGCGCTGAG 789  
DB 20 GGTGAGGATCGTCTTAAACGCTTCCCTATGATGCGCGCTATTCTCGGCGCTTGA 79  
QY 790 CGCAGCGTGTTCGCGACCATGTCATGAGCGCAGC---CTGACATCTGGGAGAAAC 846  
DB 80 CGAACCTTTTCTTGAACCAACGTCCTTCCGACCGCATCTTGAAGTGGAGGCTCAAC 139  
QY 847 GTGGCGGTGCATCGCGCGGCTGAGCTATCAAGCGCGCGCGCTGCAAGACATC 906  
DB 140 TCGACGTTACACGCGCGGAGATGAGCTCATCAACCTGTGTCAATATGTCGACGTC 199  
QY 907 GGCATTGAGCTCAAGAGATGTCGCGGATGGAGACTGTGAAGTACCGCTTCCTGGC 966  
DB 200 GCGCTGAGTTGAACAGATGTAACCGGAGAAAGAACTTCAAGTCAAGAGCTTCGCGC 259  
QY 967 TATGGCACTCTCTCGCGGCTGTCGCACTACTACGTCGCGAGCGCGCGTGAAGCTG 1026  
DB 260 TACGGGCACTCTCTCGGGTCTTATGCACTACTACGCGCGGAGCACTTGAAGTTG 319  
QY 1027 CGCAGAGACATGACACCGAGTGAAGCCCGGATGATGATCTTCATGAGCGGATGTG 1086  
DB 320 AGGAGAGACATGAGACTGTATACAGCTGTGATGCTTTCATGAGCGCAATCTC 379  
QY 1087 ATGCTGCCGAGAGGAGTCCCGGTGCGCGGCTATCGAGACACATCTTGATCT 1145  
DB 380 AGGATACAGAGAGGAGCGCTGTGAGAGGCTACAGAGAGCATGATATCATGGGTGT 438

RESULT 5  
CD296311 519 bp mRNA linear EST 16-SEP-2003  
LOCUS CD296311  
DEFINITION Strp691.007706 Sea urchin larva cDNA library MPMG691  
MPMG691IN1215;MPI\_SURUDI\_15N12 5', mRNA sequence.  
ACCESSION CD296311  
VERSION CD296311.1 GI:34747388  
KEYWORDS EST.  
SOURCE Strongylocentrotus purpuratus  
ORGANISM Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
REFERENCE 1 (bases 1 to 519)  
Poustka, A. J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,  
Reinhardt, R., Herwig, R., Panopoulou, G., and Lehrach, H.  
AUTHORS Generation, annotation, evolutionary analysis, and database  
TITLE

JOURNAL  
COMMENT

Integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)

Contact: Poustka A.J.  
Laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Imestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128

Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
(ONP) to reduce sequencing redundancy. According to the ONP  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 inner oligonucleotides are grouped into clusters. One  
clone per ONP cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
the other clones assigned to the same ONP cluster as the clone from  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA  
clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)

PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGCGAGCGGAGGAGTG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CGGTCGCGAATCCCGGT-3' pSport3/86  
High quality sequence stop: 519.  
Location/Qualifiers

## FEATURES

source

1..519  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="MPMG691IN1215;MPI\_SURUDI\_15N12"  
/cissue\_type="whole larva"  
/dev\_stage="larva 2-3 weeks"  
/lab\_host="E.coli, XLI blue"  
/clone\_lib="Sea urchin larva cDNA library MPMG691"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-TGACGATGTTAGTGGAGCGGCGCC (T)15-3' and a  
SalI 5'-TCGACCCAGCGCTCCG-3' adapters (Gibco BRL)"

## ORIGIN

Query Match 10.2%; Score 123.6; DB 6; Length 519;  
Best Local Similarity 55.9%; Pred. No. 8.9e-17;  
Matches 275; Conservative 0; Mismatches 200; Indels 9; Gaps 2;

QY 10 GACATGTTGACGATGATGAATGCAACGCGGAGAAAGTATTCGCCGTTTTCGAT 69  
DB 24 GAGATGCCAAGTTGATGATGACCTACGAAATGATGAAAGGCGGCAACCTTCCACCC 83  
QY 70 GCGGAGATGACCCCGCGGCAAAAGCACTTGGCGGTGATGAGCCAAAGCAATGCTAT 129  
DB 84 GCGGAGATGACACGCGCGGTGAGAGCGCTCCGAAAGCCATGCTGACAGGAGAAATCGAA 143  
QY 130 GCGGCGCTGTCACTCTTATCACTGATCACTACTTTCGCGGTGAGTACTAT 189  
DB 144 GCGGACTCTGAGCTCAATGACACATCAATCAATCTTCTGACTACTCTTACTGACG 203  
QY 190 TTGCGACCAAGTACGCGATGATGATGACCAACAGCGGCAAGCATTTTCGCGCGC 249  
DB 204 GTGCGTCTTCCCTTACGCGGCTTGTATCAACATGAGCAAGAGTGTCAATGCTGCGCTC 263  
QY 250 ATGAGCGGCGGCGGCGCTGCGG-----CGCAGCTTGGGAGCAACATCACTACAC 303  
DB 264 GTGCAACAGCGGCGGCGCTGCGGAGATCAACCGGTGAGCGAGTGTGATATCAACAG 323  
QY 304 GATGCGGCGGCGGCAATTTCTATGCGCGGT---GGCGGAGTACAGCGGCGGCGGAG 360  
DB 324 GATTGGCATTCGCAACATTTCTGAGGCGCGCTTGGGCACTGCTCTGGCAAGCCTTCGAG 383  
QY 361 CGCATGCGGATGATGATGACCAAGTATGCACTTTCGCGCGGCAAGTATGAGAGAGCC 420  
DB 384 AAGATCGAGGCGGATTTGATCATATACCTCATGAGAGGAAATTAAGTAAACAACT 443



QY	DB	QY	DB	RESULT 6	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
421	444	481	504	CP919042	CP919042	BFL0531_000127 Amphioxus 26 hrs CDNA library (Name convention: BFL26 or MPMG531) Branchiostoma floridae CDNA clone MPMG531.07115;BFL26_115L7 5', mRNA sequence.	CP919042	CP919042	EST.	Branchiostoma floridae (Florida lancelet) Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.	1 (bases 1 to 551)	Panopoulos,G., Hennig,S., Groch,D., Krause,A., Pousetka,A.J., Herwig,R., Vingron,M. and Lehrach,H.	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes	Genome Res. 13 (6A), 1056-1066 (2003)	12799346	Contact: Panopoulos G		

Laboratory 145, dept. Lehrstuhl  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihmestr.63-73, D-14195 Berlin, Germany  
 Tel.: +49 30 8413 1235  
 Fax: +49 30 8413 1128  
 Email: panopoul@molgen.mpg.de  
 The library was characterised by oligonucleotide fingerprinting  
 (ONFP) to reduce sequencing redundancy. According to the ONFP  
 procedure clones giving the same hybridisation pattern with a  
 battery of 200 8mer oligonucleotides are grouped into clusters. One  
 clone per cluster is selected for sequencing. The size of each  
 cluster is an indicator of the frequency of a transcript in the  
 analysed library. The cluster size as well as the coordinates of  
 all clones assigned to the same fingerprint cluster as the clone  
 from which the above EST is generated is available at the amphiXus  
 project site at <http://www.molgen.mpg.de/amphiXus>.  
 Clones and filters are distributed via the Resource Center/Primary  
 Database of the German Genome Project (<http://www.rzpdp.de>).  
 PCR Primers  
 FORWARD: 5' CCCAGGCTTACACTTATGCTCCGGCTCG 3' (M13RSP)  
 BACKWARD: 5' GCTATTCCGACGAGTCGCAAGGAGGATGTG 3' (M13FSP)  
 Insert Length: 1200 Std Error: 200.00  
 Seq primer: 5'-CCGTCGCGGAATTCGCCGGT-3' pSPORT3/86  
 High quality sequence stop: 551.  
 Location/Qualifiers  
 1..551  
 /organism="Branchiostoma floridae"  
 /mol\_type="mRNA"  
 /strain="wild type"  
 /db\_xref="taxon:7739"  
 /clone="MPMGp531i07115,BFL26\_115L7"  
 /cbase\_type="whole embryo"  
 /dev\_stage="26 hrs (neurula stage)"  
 /lab\_host="Escherichia coli, XL1 blue"  
 /clone\_lib="AmphiXus 26 hrs cDNA library (Name  
 convention: BFL26 or MPMGP531)"  
 /note="Vector: pSPORT1; Site1: SalI, KpnI, EcoRI (5');  
 Site2: NotI, BamHI, HindIII (3'); OligoDT primed and  
 directionally cloned in pSPORT vector using a NotI  
 (5'-pGACTGTTCTTACATCGCGAGCGCGCCC (7)15-3' and a SalI 5'-

Query Match	6.9%;	Score 83.2;	DB 7;	Length 551;
Best Local Similarity	57.8%;	Pred. No. 6.9e-08;		
Matches 148;	Conservative 0;	Mismatches 108;	Indels 0;	Gaps 0
Db	26	TGAATATGCAACAAGCGAGAAAAGATTATTCGCGCTTTCGATGCGCGAGATGACCCGCC	85	
Db	286	TGACATGCAATTAAGCGGACAGAGTCAACCGACTTCTCAGTGGAGAACTCCAGAGAA	345	
Qy	86	GCCAAAACGACGTTTCGGCGCTGATGCGCCAAAGAACATATGATGCGCGCGCTTTCACTT	145	
Db	346	GCGTGGACAAGCTGCGGCTCTCATGTCTACGACTAATATAGATGACGCTTCTTCACTT	405	
Qy	146	CTTATCATGTCATCACTACTATATTCGGCGCTGCTATGCTATTTGGACGCAAGTACG	205	
Db	406	CTTATCATTAATAATTAATTACTACTGCGCACTTTCTGTACATTTATTTGGTGGACCTACG	465	
Qy	206	GCATGTGTCATGCACCAACAACGCGCAGATTTGGCCGGCATGACGCGGCCAGC	265	
Db	466	GCGGTGTGTACCCATGAGCAAAAGTCTCTTCTTATCCCAAGCGGTGACGTTGTGACG	525	
Qy	266	CTGTGCGCGCGACGTT	281	
Db	526	CTGTGAGAGAGACAT	541	

RESULT	7
B1387857	
LOCUS	
DEFINITION	B1387857 553 bp mRNA linear EST 26-AUG-2003
ACCESSION	BFJL26.002560 Amphioxus 2shr cDNA library (Name convention: BFJL26
VERSION	MPMG531) Branchiostoma floridae cDNA clone MPMG531U07115 5', MRNA
KEYWORDS	sequence.
SOURCE	B1387857 B1387857.1 GI:30922696
ORGANISM	EST. Branchiostoma floridae (Florida lancelet) Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
REFERENCE	1 (bases 1 to 553)
AUTHORS	Panopoulou G., Hennig S., Groth D., Krause A., Pousetka A.J., Herrwig R., Vingron M. and Lehrach H.
TITLE	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE	22683279
PUBMED	12799346
COMMENT	Contact: Panopoulou G Laboratory 145, dept. lehrach Max-Planck-Institut fuer Molekulare Genetik Inhestr.63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: panopoul@molgen.mpg.de The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: http://www.molgen.mpg.de/amphioxus/ Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project ( <a href="http://www.rzp.d.de">http://www.rzp.d.de</a> ) PCR Primers FORWARD: 5' CCCAGCGCTTTACACTTATNGCTTCGGCCTCG 3' (M13RSP) BACKWARD: 5' GTATTATCACGCCAGCTGCAGAAAGGGGGAATGTG 3' (M13FSP) Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGGTCCGGAATTCGGCGGT-3' pSPORT3/86  
High quality sequence stop: 553.  
Location/Qualifiers

FEATURES  
source

1..553  
/organism="Branchiostoma floridae"  
/mol\_type="rRNA"  
/db\_xref="taxon:7739"  
/clone="MPMG531J07115"  
/feature\_type="whole embryo"  
/dev\_stage="26 hrs (neutula stage)"  
/lab\_host="E.coli, XL1 blue"  
/clone\_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMG531)"  
/note="Vector: pSPORT1 (Gibco BRL); Site 1: Sali, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSPORT1 vector using a NotI (5')-pGACTAGTCTAGATCGGCGCGCC (7)15-3' and a Sali 5'-TCGACCCAGCGTCCG-3' adapters (Gibco BRL)."

## ORIGIN

Query Match 6.3%; Score 83.2; DB 4; Length 553;  
Best Local Similarity 57.8%; Pred. No. 6.9e-08;  
Matches 148; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

26 TGAATGGAACAGCGGAGAAAGATTATTCGCGTTTCGATGCGGAGATACCGGCC 85  
287 TGACATGCAATACGGGAGAGAGTCAACCGACCTTCTAGTGGAGAACTCAGAGAA 346  
86 GCCAAACGACGTTCCGCGGTGATGCGCAAGACATGTCATGCGCGCTTTACCT 145  
347 GGCTGACACAGCTGCGTGTCTCATGCTGACGATATATAGATGACGCTCTTCACTT 406  
146 CTATATCATGCACTACTATTCGCGGTGCTGATGCTATTTTCGACGCAATAG 205  
407 CCTATCATTAACCACTACTACTGACCTTCTTACACTTATTCGATCGGACCTAG 466  
206 GCATGTCATTCGACCAACCAACGACGATTCGCGCGGCGATCGACGCGCGCAGC 265  
467 GGTGTGTCGTCACGACGACAAAGTGTCTTCTTATCCGACCGCTTGAAGGTGTCAGC 526  
467 GGTGTGTCGTCACGACGACAAAGTGTCTTCTTATCCGACCGCTTGAAGGTGTCAGC 526  
266 CCTGGCGCGCAGCTT 281  
527 CCTGGAGAGAGCAT 542

RESULT 8  
B2894814 524 bp DNA linear GSS 30-JUL-2003  
LOCUS B2894814  
DEFINITION Hg4\_0102 Hg pUC18 Library Halobaculum gomorrense genomic 5',  
genomic survey sequence.

ACCESSION B2894814  
VERSION B2894814.1 GI:33345290  
KEYWORDS GSS.  
SOURCE Halobaculum gomorrense  
ORGANISM Halobaculum gomorrense  
Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
Halobacteriaceae; Halobaculum.

REFERENCE  
AUTHORS Goo, Y., Roach, J., Glusman, G., Balliga, N.S., Deutsch, K., Pan, M.,  
Dasgupta, S., Ng, W. V. and Hood, L.

TITLE Low-pass sequencing for microbial comparative genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Goo Y  
Institute for Systems Biology  
1441 North 34th Street, Seattle, WA 98103, USA  
Tel: 206 732 1412  
Fax: 206 732 1299  
Email: ygoosystemsbiology.org  
Seq primer: M13 forward  
Class: shotgun.

FEATURES  
source 1..524  
Location/Qualifiers  
/organism="Halobaculum gomorrense"

/mol\_type="genomic DNA"  
/strain="ARCC 700876"  
/db\_xref="taxon:43928"  
/clone\_lib="Hg pUC18 Library"  
/note="Vector: pUC18; Site 1: SmaI; A shotgun library was  
constructed from Halobaculum gomorrense genomic DNA using  
pUC18/SmaI/BAP plasmid"

## ORIGIN

Query Match 5.7%; Score 69.6; DB 8; Length 524;  
Best Local Similarity 49.7%; Pred. No. 6.8e-05;  
Matches 177; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

764 TTTTGCTTACTACACCGCGCTGAGCGCCAGCGCTTTGCGACCATGTCATGACGCA 823  
19 TCGCGGCTACTCTCTCGTAAACGCGTGTCCGCCGACCTTCAGACGCGACGCGC 78  
824 GCTTCGATCTGAGAGAAAGACGTGCGCTGATCGCGCGGCTCGAGTATCAAGC 883  
79 AGTGGTCCGCGGAGAGAGCTCGAGGTTCCGCCGACCGGACCGGAGTGTCAAGC 138  
884 CGGCGCGCGCTGCAAGACATCGCATCGAGCTCAAGAGATTAACCGGAGTGGAC 943  
139 CCGACGAGGTGACGACCGCGCAGACCTCGGATCTACGCGGAGGTCAACGCGAGCGCC 198  
944 TGCTGAATGACCGCTCTTCCGCTATGCGCACTCTTCGCGGTGCTGCGCACTACTAG 1003  
199 TGCAAGACTCTTCGACCGACCAACTCATCTTTCGCGTGAAGATGCTCTTCTGCA 258  
1004 GTGCGAGCGCGCGCTGAGGTGCGCGAGACATGACGACGACGAGCTGAAGCCGCGATG 1063  
259 GCCAGCGCTTACAGCTGAAGCGCGGCGCATCATCTTACCGGAGCGCGCGCGCTG 318  
1064 TGGTTCATGAGGCGATGCTGATGCTGCGGAGGCGCATGCGCGTGGCGCGC 1119  
319 GCGTTCACCGGAGCGCGCGGTGCTGCGTGAAGGCGACCGCTCATCTATGCGC 374

RESULT 9  
CB656525 722 bp mRNA linear EST 09-APR-2003  
LOCUS CB656525  
DEFINITION OSJNEC10P11.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEC10P11 5', mRNA sequence.

ACCESSION CB656525  
VERSION CB656525.1 GI:29660250  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriatroidae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,  
Kudrna, D., Dean, R., Soderlund, C., Wang, R. and Wang, G.  
large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
COMMENT Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aac cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g

Plate: 10 row: P column: 11  
Seq primer: gta aac cga cgg cca gtc.

FEATURES  
source 1..722  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"



DB	692	AAGGTGTCGGCAGCTTCAAGACCGGCGACATCACCGG	719
RESULT 11			
LOCUS	CNS0091P/c		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR1916 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL053013		
VERSION	AL053013.1		
KEYWORDS	GS...		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 925)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)		
COMMENT	- Web : www.genoscope.cns.fr Determination of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> . location/Qualifiers		
FEATURES			
source	1..925		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/clone_1ib="BACR1916"		
	/clone_1ib="RPCI-98"		
	/note="end : TET3"		
ORIGIN			
Query Match	5.2%	Score 63.6;	DB 9; Length 925;
Best Local Similarity	15.5%;	Pred. No. 0.0015;	
Matches	55;	Conservative 171;	Mismatches 125; Indels 3; Gaps 1;
QY	228	CGCCACGACGATTTCCGCCGCGATCGACGCGCGACGCCCTCGCCACCTTCGGCA	287
DB	915	CSBSBCSSSMSTSSNSBSCSSGSSSSSTSSMSSBSBSSGSSSSSGTSSACVX	856
QY	288	CAACATCACTCAACCGCATGCGCGCGGACATATTTATGCGCGCTGCGCCAGCTGAC	347
DB	855	CNASSGCCCGCGMAABCCMCSSSSCCSASABKVAASGAGKRGGSAGASASHSS	796
QY	348	CAGCGGCGCCAGCGCATCGCATCGCATGCAATTCGACACCGTCATATTCGATTCGGCGCA	407
DB	795	SACBSSSSSCASCMASBSSSSASBSRSGGAGAGGASBSRSSSSSASAGVSSA	736
QY	408	GCTGAGAGAACCCCTA---CCGGCGTGCATCTGTCGACATCAAGCCCTCGATGTG	464
DB	735	SSSSSSCGSVSVCSVSAWSBSCSSBSSSASASSSSSSSSASCSACCTTWSGSCSTS	676
QY	465	GATCGCGCACATCAAGTCTCGAAGACAGAGCTATCGCGAAGCGCCCGGTGTG	524
DB	675	ASMAAABSSSSSSSSSSSSSABSSASSSSSSSSSSSSSSACBBSMSSGCGSGS	616
QY	525	TGACGTGCGCGCGCGCGCTCGCGCGTGCATCAAGCGGCGGTGCCGAGCA	578

[illegible]

Db 527 GTGACGTCTTCGACCGGTGCGGGCCGACGAAGCGCGGAGCGTCTCGC 586  
 Qy 1030 GAGACATCGACACCGAGTGAAGCCCGCATGCTGTCTTCATGAGCCGATGCGATG 1089  
 Db 587 GGGAAAGACGACCGCCCGCTGGCGCGCGCGCATGTCTTCGCAAGCGGACGTGAGTGC 646  
 Qy 1090 CTGCGGAGGAGCATGCGCGGTGCGCGG 1117  
 Db 647 CTGACGACGAGCTCGCCGCTACGACG 674

RESULT 13  
 CB659172 812 bp mRNA linear EST 09-APR-2003  
 LOCUS OSJNEC15018.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA  
 DEFINITION clone OSJNEC15018 5', mRNA sequence.  
 ACCESSION CB659172  
 VERSION CB659172.1 GI:29662897  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 1 (bases 1 to 812)  
 Jantsuiliyarc,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea  
 Unpublished (2003)  
 JOURNAL Contact: Rod Wing  
 COMMENT Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR PRIMER  
 FORWARD: gta aac cga cgg cca gtc  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 15 row: 0 column: 18  
 Seg primer: gta aac cga cgg cca gtc.

FEATURES  
 source  
 1..812  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNEC15018"  
 /issue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEC"  
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN  
 Query Match 5.0%; Score 60.8; DB 6; Length 812;  
 Best Local Similarity 47.2%; Pred. No. 0.006; Mismatches 207; Indels 0; Gaps 0;  
 Matches 185; Conservative 0;

Qy 511 GGGCGCCGCGTGTGTGATCGTGGCGCGGCGCTGCGCGGCTCCATCAAGCGCGCGT 570  
 Db 255 GGAAGAGGCGCGCGCGTGGCGGTGCGGAGCTGGCGCGATCTTTGAGAGCGGCGT 314  
 Qy 571 CCGGACATGAAATGGCGATCGCACCAACCAATGCGATGATCCGCGAGATCGCAATG 630  
 Db 315 GCGCGCGCGCGCTGTGCGCTGCGCAATCAAGATGCGCGCGCGCGCGCGCGTGG 374  
 Qy 631 TTCCCTTCGTGAGTGTGATGACACTGACTGTTCAGTGGGCAATCAACACGAC 690

Db 375 TCGAGTTGACAGTGTCTGAACGCGCGCGCGAGAGACATGATGATGACAGCC 434  
 Qy 691 GGGCGCCAAATTCGGTCAACCGCATGTGCAATCCGGGACATCTTTGGCTCAAC 750  
 Db 435 GCGGTCTAGCGCGCTGACCAAGTCCGACGTGAACGGCTCTTCGCGACCTCGAGTCC 494  
 Qy 751 ACCCTCCGATGATCTTGGGCTACTACACCGCGCTGAGACGACGCTGTTCTGCAACAT 810  
 Db 495 CGGAAGATCATGATCAATCACTTCACTACCGTGGCAACCTCAAGCAAGCTGACGCG 554  
 Qy 811 GTGATGACGCCAGCTCGACATCTGGAGAAACGTGCGCTGCATGCGCGGAGTTC 870  
 Db 555 TTGAGTACGCGGACCAAGTCTGGAGAGTGTATCCGACGCTGAGAGAGGCAACAC 614  
 Qy 871 GAGCTGATCAAGCCGCGCGCGCTGCAAGGA 902  
 Db 615 AAGGCGATGTGTGTCTGTCAACAGCCAGAA 646

RESULT 14  
 AUI62766 604 bp mRNA linear EST 03-APR-2002  
 LOCUS AUI62766 Rice mature leaf Oryza sativa (japonica cultivar-group)  
 DEFINITION cDNA clone S21656, mRNA sequence.  
 ACCESSION AUI62766  
 VERSION AUI62766.1 GI:11026165  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 1 (bases 1 to 604)  
 Sasaki,T. and Yamamoto,K.  
 Rice cDNA from mature leaf (2000)  
 Unpublished (2000)  
 JOURNAL Contact: Takuji Sasaki  
 COMMENT National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abr.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/

FEATURES  
 source  
 1..604  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="S21656"  
 /issue\_type="mature leaf"  
 /clone\_lib="Rice mature leaf"

ORIGIN  
 Query Match 5.0%; Score 60.6; DB 1; Length 604;  
 Best Local Similarity 44.8%; Pred. No. 0.005; Mismatches 285; Indels 0; Gaps 0;  
 Matches 231; Conservative 0;

Qy 634 CCCTTCGTGAGCTGTGATGACACCTGAGCTGTTCAGTGGGCAATCAACCGAGCG 693  
 Db 57 CGCCCATCTGTGTAAAGAGAGTGTGCGGAGATGACCGCTCTCTCACTTCGCC 116  
 Qy 694 GGGCAAAATCCGATCAACCAACCGATGTGCAATCGGAGACATCTTTCGCTGAACCC 753  
 Db 117 GGGCGACCGCGCGGACCTTACCAATGAGTCTGCGCTCCACCGCTTCGTCGATC 176  
 Qy 754 TTCCGATGATCTTGGGCTACTACACCGGCTGAGAGCGACGCTGTTCTGCAACATG 813  
 Db 177 CTCTACTGATGGGGGCTTCGACCGCGCTTGGCGCTTCTCTCTCTCTCTCTGTC 236  
 Qy 814 GATGACGCGAGCTGACATCTGGAGAAAGATGAGCGTGCATGCGCGGCGCTGAG 873

Db 237 GTCTCTACAGCGCTCTCTGTCGCCAGAGCTTGGGCTCGCATGGCGCGCTGCTCATG 236  
 QY 874 CTGATCAAGCCGGGCGCGCTGTCAGAGCATGCGCCATGCAAGATGTACCGC 933  
 Db 297 GAGCTCAAGGAGGACACACCTCTGCTCGTCATCAACATGGTCTTCTCATGCGCGG 356  
 QY 934 GATGTGGAGCTGTGAAGTACCGCTCTTGGCTATAGCCATCTCTGGCGCTGTCTGC 993  
 Db 357 GGCTACTACGTCCAGCATCTCCGCTCTGTGGCTGTGGCTGTGCTCACTACAGC 416  
 QY 994 CACTACTACGCTGCGAGCGCGGCTGTGAGCTGTGCGAGCAATCAACCGAGCTGAAG 1053  
 Db 417 TTCTTACTGCTACCGCTCTCTCATGCGCATCAAGTTCGGCGACGCGCGCTCACTAGAC 476  
 QY 1054 CCGCGCATGTGTGTCTTCATGAGCCGATGTGTATGCTGCGGAGGCGCATGCCGCTGCC 1113  
 Db 477 TGGGCGGCGCGGCGCGCGCTGCTGCTGCGCGACTTCCCGGCAATCAAGGCGCTGGG 536  
 QY 1114 GGGCGCTATGCGAGCAGCATCTCTGATGCTGGG 1149  
 Db 537 ATCAACAACCATGTGGTGAAGTCTGCGTATGANG 572

## RESULT 15

CD225047 668 bp mRNA linear EST 21-MAY-2003  
 LOCUS CCCL\_37\_H01.g1\_A007 Callus culture/cell suspension Sorghum bicolor  
 DEFINITION cDNA clone CCCL\_37\_H01\_A007 5', mRNA sequence.

ACCESSION CD225047  
 VERSION CD225047.1 GI:30968481

KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor

REFERENCE Sorghum bicolor  
 Sukaryoca; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

1 (bases 1 to 668) Cordamer-Pratt, W.-M., Wentzel, V., Suzuki, Y., Sugano, S.,  
 Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Rathore, K.,  
 Eastman, A. and Pratt, L.H.

An EST database from Sorghum: callus culture and cell suspension  
 Unpublished (2003)

Other\_ESTs: CCCL\_37\_H01.b1\_A007

Contact: Cordamer-Pratt WM

Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;

sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.

Seq primer: Sug5 (CTTCTGCTTAAAGTCGC).

location/Qualifiers

1..668

/organism="Sorghum bicolor"

/mol\_type="mRNA"

/cultivar="RTx430"

/db\_xref="taxon:4558"

/clone="CCCL\_37\_H01\_A007"

/lab\_host="DH10B-T1 phage-resistant E. coli"

/clone\_lib="Callus culture/cell suspension"

/note="Vector: pME185-FL3; Site 1: XhoI; Site 2: XhoI; The  
 library was prepared from a mixture of polyA+ RNA from  
 callus culture tissue and cells in suspension culture.  
 Double-stranded cDNA was cloned unidirectionally into  
 different DraIII sites of the pME185-FL3 vector (5-prime

DraIII site is CACTGTGTG, 3-prime DraIII site is  
 CACCATGTG). XhoI excises the cDNA insert."

ORIGIN  
 Query Match 5.0%; Score 60.6; DB 6; Length 668;  
 Best Local Similarity 47.5%; Pred. No. 0.0066;  
 Matches 180; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 171 CGGCTGGCTGTACTGCTATTGAGCGCAAGTACGGCATGTGATGACCAACAAGC 230  
 Db 285 CCGCAAGTTCACCAACGCTGTGCGCAACAGGACCAAGCCCACTACAGATCACTG 344  
 QY 231 CAGAGATTTGGGCGCGGATGACGCGCGCCAGCCCTGGCGGAGCTTGGCGACA 290  
 Db 345 CACCAACGACTGGGGGCTCTGCTACATCAAGTCCCTGCTGCTGCCCCAAGTG 404  
 QY 291 CATCACTTACACGCACTGGCGCGGCAATTTCTATCGGCGTGGCGGACGTCAC 350  
 Db 405 CTTGCTTACTGTGGCTTACTGCTTCACTTGTGATGTGATGCGGGGCACTC 464  
 QY 351 GGGCGCAAGCGCATGCGATGAGTTGACCAAGTCAATCTGACTTCCGCCAGCT 410  
 Db 465 GTGCGGCGACCGCGGTTACGCGCGCGCGCAACACTTCTACTTCCACGCGAAG 524  
 QY 411 CGAGAGGCTTACCGGGCTGCACTTCTGCAATCAACCGCCCTCGATGTGATGG 470  
 Db 525 GAGCAGAGCTTCTGCTGCTGCTCGACGACCGCTTCAATCAAGCGCGTTCAAGG 584  
 QY 471 CACCATCAAGTGTGCTGGAAGACAGAAAGTGAATCCGGAAGGCGCGGCTGTGACGT 530  
 Db 585 CAACCAACAGCGCGACTCGGCGCGGAGCTTCACTGAGTGAAGGCGCTGGGCTCACTT 644  
 QY 531 CGGCGCGCGCGCTGCGG 549  
 Db 645 CAGCGCGCGCGAGCGGCGG 663

Search completed: November 19, 2004, 19:22:01  
 Job time : 4241 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 12:40:45 ; Search time 653 Seconds

(without alignments)  
9743.183 Million cell updates/sec

Title: US-10-807-228-2

Perfect score: 1212  
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACATCATCCGCAAC 1212

Scoring table: IDENTITY\_NUC  
Gapop 10.0' , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## ALIGNMENTS

Database : N\_Geneseq\_23Sep04:\*

1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	2	AAT61367 Creatinin
2	1210.4	99.9	1212	2	AAT38807 Creatine
3	1210.4	99.9	1212	2	AAV35699 Stable cr
4	1210.4	99.9	1212	4	AA814742 Creatine
5	1210.4	99.9	1212	6	AA199856 Alkaligen
6	1210.4	99.9	1212	6	ABA93696 Creatine
7	1184.8	97.8	1215	8	AA113291 Creatine
8	981.6	81.0	1215	8	ACC69514 Erwinia s
9	981.6	81.0	1215	8	ACC69519 Mutant Er
10	978.4	80.7	1215	8	ACC69515 Mutant Er
11	976.8	80.6	1212	8	ACC69517 Mutant Er
12	976.8	80.6	1212	8	ACC69516 Mutant Er
13	976.8	80.6	1215	8	ACC69518 Mutant Er
14	975.2	80.5	1212	8	ACC69521 Mutant Er
15	973.6	80.3	1212	8	ACC69520 Mutant Er
16	640.2	52.8	1209	2	AA105502 Thermos
17	635.6	52.4	1212	1	AA60274 Sequence
18	634	52.3	1212	1	AA81271 Sequence
19	563	46.5	1134	1	AA170877 Sequence
20	456.2	37.6	1233	1	AA190635 Sequence
21	317.6	26.2	1282	2	AAV66090 Arthrobac

22	74	6.1	858	11	ABD05567	ABD05567 Pseudomon
23	74	6.1	963	11	ABD05831	ABD05831 Pseudomon
24	74	6.1	1176	11	ABD05710	ABD05710 Pseudomon
25	72.4	6.0	786	10	ADG73350	ADG73350 Pseudomon
26	72.4	6.0	786	10	ADG73352	ADG73352 Pseudomon
27	69.6	5.7	985	6	ABO44816	ABO44816 Oligonuc1
28	69.6	5.7	985	6	ABO44817	ABO44817 Oligonuc1
29	68.8	5.7	65140	4	AA117184	AA117184 Streptomy
30	68.8	5.7	125401	4	AA117186	AA117186 Streptomy
31	67.8	5.6	1227	6	ABL61294	ABL61294 N. unifor
32	66.8	5.5	2712	8	ACA36793	ACA36793 Prokaryot
33	66.8	5.5	110000	6	ABA03041_18	Continuation (19 o
34	66	5.4	2037	8	ACC44573	ACC44573 Glucoamyl
35	65.2	5.4	690	3	AAA50948	AAA50948 Humanised
36	65	5.4	2052	8	ACA23861	ACA23861 Prokaryot
37	64.8	5.3	3157	2	AAV08896	AAV08896 Cellobios
38	64	5.3	2364	8	ABV74978	ABV74978 Synthetic
39	63.6	5.2	690	6	AAD46285	AAD46285 Anemonia
40	63.6	5.2	2064	2	AAO52638	AAO52638 Streptomy
41	63.6	5.2	9546	12	ADN40876	ADN40876 Plasmid p
42	63.4	5.2	1083	12	ADJ39271	ADJ39271 Plant CDN
43	62.8	5.2	9785	6	ABK91624	ABK91624 Modified
44	62.4	5.1	1294	2	AAZ06824	AAZ06824 Streptomy
45	62.4	5.1	1294	4	AAH74537	AAH74537 Nucleotid

## RESULT 1

AAT61367 standard; DNA; 1212 BP.

AAT61367;

17-APR-1997 (first entry)

Creatinine amidinohydrolyase coding sequence.

Thermal stability; creatinine amidinohydrolyase; creatine; sarcosine;

urea; blood; ds.

Alcaligenes faecalis.

JP08308579-A.

26-NOV-1996.

16-MAY-1995; 95JP-00117283.

16-MAY-1995; 95JP-00117283.

(TOYM ) TOYORO KK.

WPI; 1997-059698/06.

P-PSDB; AAM11861.

Gene coding for creatinine amidinohydrolyase - used to quantify blood or

urinary creatinine as a disease indicator.

Claim 4; Page 10-11; 12pp; Japanese.

This sequence encodes a thermally stable creatinine amidinohydrolyase

which has a low Km value for creatine. The creatinine amidinohydrolyase

has the following physicochemical properties: (a) action: converts

creatine and water into sarcosine and urea; (b) optimum temp.: 40-45

deg.C; (c) optimum pH: 8.0-9.0; (d) thermal stability: about 50 or less

(pH 7.5, 30 mins); (e) stable pH: a pH range of about 4-10; (f) Km value:

about 15.2 mM; (g) mol. wt.: about 4300 (SDS-PAGE); and (h) isoelectric

point: about 3.5. The creatinine amidinohydrolyase may be used in the

quantification of blood or urinary creatinine and creatine as an

indicator of various diseases



CC uremia, chronic nephritis, gigantism and tonic muscular dystrophy. The  
 CC enzyme has a lower km value than prior art creatine amidinohydrolase (cf.  
 CC US 5451520). (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 2; Length 1212;

Best Local Similarity 99.9%; Pred. No. 1.5e-195; Mismatches 1; Indels 0; Gaps 0;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTCGATGTAATGCAACAACGCGGAGAAAGATTATTCGCG 60  
 DB 1 ATGACTGACGACATGTCGATGTAATGCAACAACGCGGAGAAAGATTATTCGCG 60  
 QY 61 TTTTGGATGCCGAGATGACCCGCCCAAAGACGTTCCGGCTGGATGGCCAAAG 120  
 DB 61 TTTTGGATGCCGAGATGACCCGCCCAAAGACGTTCCGGCTGGATGGCCAAAG 120  
 QY 121 AATGTCGATGGGGGCTGTCACCTTATCACTGATCACTAATTCGCGGCTGG 180  
 DB 121 AATGTCGATGGGGGCTGTCACCTTATCACTGATCACTAATTCGCGGCTGG 180  
 QY 181 TACTGCTATTTCGACGCAAGTACGCGCATGTCATGACCAACAACGCGCAGATT 240  
 DB 181 TACTGCTATTTCGACGCAAGTACGCGCATGTCATGACCAACAACGCGCAGATT 240  
 QY 241 TCGGCGGCGCATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 DB 241 TCGGCGGCGCATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 QY 301 ACCGATGCGCGCGCGGAGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCG 360  
 DB 301 ACCGATGCGCGCGCGGAGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCG 360  
 QY 361 CGCATGCGCGCATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
 DB 361 CGCATGCGCGCATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
 QY 421 CTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480  
 DB 421 CTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480  
 QY 481 TCGGTCGGAAGAGCAAGCTGATCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540  
 DB 481 TCGGTCGGAAGAGCAAGCTGATCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540  
 QY 541 GCGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
 DB 541 GCGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
 QY 601 AATGCGATGATCCGCGGAGATGCGCAATGTTCCCTTCTGTCGAGCTGATGCA 660  
 DB 601 AATGCGATGATCCGCGGAGATGCGCAATGTTCCCTTCTGTCGAGCTGATGCA 660  
 QY 661 ACCGTCGTCAGTCGCGGCGATCAACCGCGCGGCGGCGGCGGCGGCGGCGG 720  
 DB 661 ACCGTCGTCAGTCGCGGCGATCAACCGCGCGGCGGCGGCGGCGGCGGCGG 720  
 QY 721 GTGCAATCGCGGAGATCCTTTGCTCAACCTTCCCGGATGATCTTGGGCTACTA 780  
 DB 721 GTGCAATCGCGGAGATCCTTTGCTCAACCTTCCCGGATGATCTTGGGCTACTA 780  
 QY 781 GCGCTGAGCGCAACGCTGTTCTGACGATGTCATGACGCGGCGGCGGCGGCGG 840  
 DB 781 GCGCTGAGCGCAACGCTGTTCTGACGATGTCATGACGCGGCGGCGGCGGCGG 840  
 QY 841 AAGAACGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 DB 841 AAGAACGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 QY 901 GACATCGCGCATGAGCTCAACGAGATGTAACGCGGAGTGGGACCTGTAAGTAC 960  
 DB 901 GACATCGCGCATGAGCTCAACGAGATGTAACGCGGAGTGGGACCTGTAAGTAC 960

QY 961 TTGCGCTATGCGCACTCCTTTGGGCGTGTGTCGCACTACTACGCGGAGCGGCGG 1020  
 DB 961 TTGCGCTATGCGCACTCCTTTGGGCGTGTGTCGCACTACTACGCGGAGCGGCGG 1020  
 QY 1021 GAGCTGCGCGAGGACATGACACCGAGCTGAAGCCCGGATGTCGTCCTCAATGAGCG 1080  
 DB 1021 GAGCTGCGCGAGGACATGACACCGAGCTGAAGCCCGGATGTCGTCCTCAATGAGCG 1080  
 QY 1081 ATGTCGATGCTGCGGAGGCGGATCCCGGTCGCGGCGGCTATTCGCGACAGCATCTG 1140  
 DB 1081 ATGTCGATGCTGCGGAGGCGGATCCCGGTCGCGGCGGCTATTCGCGACAGCATCTG 1140  
 QY 1141 ATGTCGCGGAGGAGCGGTCGCGGAGATCAACCGGCTTCCGTCGTCGCGGAGCACAC 1200  
 DB 1141 ATGTCGCGGAGGAGCGGTCGCGGAGATCAACCGGCTTCCGTCGTCGCGGAGCACAC 1200  
 QY 1201 ATCATCCGCAAC 1212  
 DB 1201 ATCATCCGCAAC 1212

# RESULT 3

AAV35699  
 ID AAV35699 standard; DNA; 1212 BP.

AAV35699;

16-SEP-1998 (first entry)

Stable creatine amidinohydrazide encoding DNA.

Creatine amidinohydrazide; mutant; stable; enzyme; diagnostic agent; ds.

Alcaligenes faecalis.

Key Location/Qualifiers

FT CDS 1..1212

FT /\*tag= a /transl\_except= (pos:433..435, aa:Glu)

FT /product= "Stable creatine amidinohydrazide"

FT /note= "the stop codon is not indicated"

PN JP10174585-A.

PD 30-JUN-1998.

PF 17-DEC-1996; 96JP-00337027.

PR 17-DEC-1996; 96JP-00337027.

PA (TOYM ) TOYORO KK.

DR WPI: 1998-421167/36.

DR P-PSDB: AAM61905.

PT New creatine amidino-hydrazide used as diagnostic agent - is more stable in neutral buffer than wild type creatine amidino-hydrazide.

PS Disclosure; Page 11-13; 14pp; Japanese.

CC This DNA encodes a stable creatine amidinohydrazide which is a mutant

CC creatine amidinohydrazide and has improved long-term stability in a neutral

CC buffer compared to wild type creatine amidinohydrazide. A recombinant

CC plasmid containing the stable creatine amidinohydrazide gene can be used to

CC transform a cell for the recombinant production of the enzyme. This

CC stable creatine amidinohydrazide is useful as a diagnostic agent can be

CC produced commercially

Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 2; Length 1212;

Best Local Similarity 99.9%; Pred. No. 1.5e-195;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ATGACTGACGACATGTTGTCAGCTGTATGAAATGGCAACAACGGGGAAGAAATTTATTCGCCG 60
DB 1 ATGACTGACGACATGTTGTCAGCTGTATGAAATGGCAACAACGGGGAAGAAATTTATTCGCCG 60
OY 61 TTTTGGATGCGGAGATGACCCCGCGCCAAAACGAGCTTCCGGGCTGGATGGCAAGAC 120
DB 61 TTTTGGATGCGGAGATGACCCCGCGCCAAAACGAGCTTCCGGGCTGGATGGCAAGAC 120
OY 121 AATGTGATGCGGCGCTGTTCACTTATCACTGATCACTAATTAATTCGGCTGGTG 180
DB 121 AATGTGATGCGGCGCTGTTCACTTATCACTGATCACTAATTAATTCGGCTGGTG 180
OY 181 TACTGCTATTTTCGGACGCGAAGTACGGCAATGATCAGCAACAAGCCGACGACGATT 240
DB 181 TACTGCTATTTTCGGACGCGAAGTACGGCAATGATCAGCAACAAGCCGACGACGATT 240
OY 241 TCGGCGGCGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 TCGGCGGCGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
OY 301 ACCGACTGCGCGCGGCAATTTTATGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 301 ACCGACTGCGCGCGGCAATTTTATGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 360
OY 361 CGCATCGGCGATGAGTTGCAACAAGTCAATCTGACCTTCCGCGCGGCGGCGGCGGCGG 420
DB 361 CGCATCGGCGATGAGTTGCAACAAGTCAATCTGACCTTCCGCGCGGCGGCGGCGGCGG 420
OY 421 CTACCGGCGGCGTGCACCTTGTGACATCAAGCCGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 CTACCGGCGGCGTGCACCTTGTGACATCAAGCCGCGGCGGCGGCGGCGGCGGCGGCGG 480
OY 481 TCGCTCGAAGACAGACAGCTGATCCGCAAGGCGCGCGGCGGCGGCGGCGGCGGCGG 540
DB 481 TCGCTCGAAGACAGACAGCTGATCCGCAAGGCGCGCGGCGGCGGCGGCGGCGGCGG 540
OY 541 GCGTGGCGGCGTGCATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 541 GCGTGGCGGCGTGCATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
OY 601 AATGCGATGATCCGCGGAGATGCGCAATGCTTCCCTTCCGCGGCGGCGGCGGCGGCGG 660
DB 601 AATGCGATGATCCGCGGAGATGCGCAATGCTTCCCTTCCGCGGCGGCGGCGGCGGCGG 660
OY 661 ACTTGGTTCAGTCCGCGGATCAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 ACTTGGTTCAGTCCGCGGATCAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
OY 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACCTTCCGCGGCGGCGGCGGCGGCGGCGG 780
DB 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACCTTCCGCGGCGGCGGCGGCGGCGGCGG 780
OY 781 GCGCTGAGGCGGCGGCTGTTTCGACCATGTCATGACGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 GCGCTGAGGCGGCGGCTGTTTCGACCATGTCATGACGCGGCGGCGGCGGCGGCGGCGG 840
OY 841 AAGAACGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 841 AAGAACGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
OY 901 GACATCGGCGATGAGCTCAAGAGATGTAACCGGAGTGGAGACTTCTGAATCCGCTCC 960
DB 901 GACATCGGCGATGAGCTCAAGAGATGTAACCGGAGTGGAGACTTCTGAATCCGCTCC 960
OY 961 TTCGCGTATGCGGCGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 961 TTCGCGTATGCGGCGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
OY 1021 GAGCTGCGGAGGAGCATGACACCGAGCTGAAGCCCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 GAGCTGCGGAGGAGCATGACACCGAGCTGAAGCCCGGCGGCGGCGGCGGCGGCGGCGG 1080

```

```

OY 1081 ATGCTGATGCTGCGGAGGCGGATGCGCGGCGGCGGCGGCTATCCGAGCAGACATCCTG 1140
DB 1081 ATGCTGATGCTGCGGAGGCGGATGCGCGGCGGCGGCGGCTATCCGAGCAGACATCCTG 1140
OY 1141 ATCGTGGGAGGAGCGGTGCGGAGACATCACCGGCTTCCGTTCCGTCCGGAACACACC 1200
DB 1141 ATCGTGGGAGGAGCGGTGCGGAGACATCACCGGCTTCCGTTCCGTCCGGAACACACC 1200
OY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

```

RESULT 4  
 AAS14742  
 ID AAS14742 standard; DNA; 1212 BP.  
 XX  
 AC AAS14742;  
 XX  
 DT 19-DEC-2001 (first entry)  
 XX  
 DE Creatine amidinohydrolyase genomic DNA.  
 XX  
 KW Creatine amidinohydrolyase; water; sarcosine; urea; creatinine; uraemia;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; giantism;  
 KW pigment absorbance; ds.  
 XX  
 OS Alcaligenes faecalis.  
 XX  
 FH Key  
 FT 1. -1212  
 FT CDS  
 FT  
 FT /cag= a  
 FT /product= "A. faecalis creatine amidinohydrolyase"  
 FT /transl\_except= (pos:433..435, aa:Glu)  
 FT /partial  
 FT /note= "No stop codon"  
 XX  
 EN EPI132467-A2.  
 XX  
 PD 12-SEP-2001.  
 XX  
 PF 13-FEB-1997; 2001EP-00113052.  
 XX  
 PR 13-FEB-1996; 96JP-00025435.  
 PR 13-FEB-1997; 97EP-00102270.  
 XX  
 PA (TOYM ) TOYO BOSEKI KK.  
 XX  
 PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;  
 XX  
 DR WPI; 2001-612481/71.  
 DR P-PSDB; AAU08727.  
 XX  
 PT New creatine amidinohydrolyase, useful as a routine reagent for clinical  
 PT tests for determining creatine and creatinine in biological samples,  
 PT particularly useful in diagnosing diseases such as uremia or chronic  
 PT nephritis.  
 XX  
 PS  
 PS Disclosure; Page 15-16; 21pp; English.  
 XX  
 CC The invention relates to Alcaligenes faecalis creatine amidinohydrolyase,  
 CC which catalyzes creatine and water to sarcosine and urea. Creatine  
 CC amidinohydrolyase can be produced by culturing a microorganism producing  
 CC the protein in a nutrient medium and recovering the protein from the  
 CC resulting culture. Creatine amidinohydrolyase is useful as a routine  
 CC reagent for clinical tests for determining creatine and creatinine in  
 CC biological samples. This is particularly useful in diagnosing diseases  
 CC such as uraemia, chronic nephritis, acute nephritis, giantism and tonic  
 CC muscular dystrophy. The presence of creatine in a sample can be  
 CC determined by measuring an absorbance of a pigment produced by the  
 CC reaction of a reagent containing creatine amidinohydrolyase with the  
 CC sample. This sequence represents genomic DNA encoding Alcaligenes

CC faecalis creatine amidinohydrolase

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 4; Length 1212;

Best Local Similarity 99.9%; Pred. No. 1.5e-195;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 ATGATGACGACATGTTGACGATGTAATGACAAACGGGAGAAAATTATTCGCCG 60
Db 1 ATGATGACGACATGTTGACGATGTAATGACAAACGGGAGAAAATTATTCGCCG 60
Oy 61 TTTTGGATGCGGATATGACCCCGCCAAAACGATTCGGGGTGGATGGCCAAAGAC 120
Db 61 TTTTGGATGCGGATATGACCCCGCCAAAACGATTCGGGGTGGATGGCCAAAGAC 120
Oy 121 AATGTCGATGCGGCGCTGTTACCTCTTATCACTGATCAATCACTATTCGGGCTG 180
Db 121 AATGTCGATGCGGCGCTGTTACCTCTTATCACTGATCACTATTCGGGCTG 180
Oy 181 TACTGCTATTTGACGCAAGTACGGCATGGTCAATGACCAACAGCCGACGAGATT 240
Db 181 TACTGCTATTTGACGCAAGTACGGCATGGTCAATGACCAACAGCCGACGAGATT 240
Oy 241 TCGGCGGCGATGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 TCGGCGGCGATGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Oy 301 ACCGACTGGCGCGGCGGCAATTTCTATGCGCGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 ACCGACTGGCGCGGCGGCAATTTCTATGCGCGCGGCGGCGGCGGCGGCGGCGG 360
Oy 361 CGCATCGGCATGACGATGACGACGATCAATCTGACTTCGCGCGGCGGCGGCGG 420
Db 361 CGCATCGGCATGACGATGACGACGATCAATCTGACTTCGCGCGGCGGCGGCGG 420
Oy 421 CTACCGGCGGCTGACCTTGTGACATCAAGCCGCGGCGGCGGCGGCGGCGGCGG 480
Db 421 CTACCGGCGGCTGACCTTGTGACATCAAGCCGCGGCGGCGGCGGCGGCGGCGG 480
Oy 481 TCGCTGCAAGAGCAGAGCTGATCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540
Db 481 TCGCTGCAAGAGCAGAGCTGATCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540
Oy 541 GCGTGGCGGCGTCCCATGACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 GCGTGGCGGCGTCCCATGACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Oy 601 AATGCGATGATCGCGGAGATCGCAATCGTTCCCTTGTGAGACTGATGACCTGG 660
Db 601 AATGCGATGATCGCGGAGATCGCAATCGTTCCCTTGTGAGACTGATGACCTGG 660
Oy 661 ACGTGGTTCCAGTGGGCGGATCAACCGAGCGGCGGCGGCGGCGGCGGCGG 720
Db 661 ACGTGGTTCCAGTGGGCGGATCAACCGAGCGGCGGCGGCGGCGGCGGCGG 720
Oy 721 GTGCAATCGGCGGATCTTTCGCTCAACACTTCCGATGATTCGGCTACTACACC 780
Db 721 GTGCAATCGGCGGATCTTTCGCTCAACACTTCCGATGATTCGGCTACTACACC 780
Oy 781 GCGCTGAGCGGACGCTGTTCTGCAACCAATGTCGATGACGCGGCGGCGGCGG 840
Db 781 GCGCTGAGCGGACGCTGTTCTGCAACCAATGTCGATGACGCGGCGGCGGCGG 840
Oy 841 AAGAACGTCGCGGCGGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 AAGAACGTCGCGGCGGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Oy 901 GACATCGCATGACGCTCAACGAGATGTAACGCGAGTGGGACCTCTGAAGTACG 960
Db 901 GACATCGCATGACGCTCAACGAGATGTAACGCGAGTGGGACCTCTGAAGTACG 960
Oy 961 TTGGGCTATGCGCACTCTTGGGCGTGTGTCGCACTACTACGATGCGGCGGCGG 1020

```

```

Db 961 TTGGGCTATGCGCACTCTTGGGCGTGTGTCGCACTACTACGATGCGGCGGCGG 1020
Oy 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGGCGATGTTGTCATGAGGCG 1080
Db 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGGCGATGTTGTCATGAGGCG 1080
Oy 1081 ATGCTGATGCTGCGGAGGAGGATGCGCGGTCGCGGCGGCTATCGGACACGACTCT 1140
Db 1081 ATGCTGATGCTGCGGAGGAGGATGCGCGGTCGCGGCGGCTATCGGACACGACTCT 1140
Oy 1141 ATCGTCGGGAGGAGCGGTCGCGGAGAAATGACCGGCTTCGCTCGGCGGAGACAC 1200
Db 1141 ATCGTCGGGAGGAGCGGTCGCGGAGAAATGACCGGCTTCGCTCGGCGGAGACAC 1200
Oy 1201 ATCATTCGCGCAC 1212
Db 1201 ATCATTCGCGCAC 1212

```

# RESULT 5

AA199856  
ID AA199856 standard; DNA; 1212 BP.

AC AA199856;  
XX 07-AUG-2003 (revised)  
DT 28-JAN-2002 (first entry)

XX Alkaligenes faecalis creatineamidinohydrolase encoding DNA.

XX Alkaligenes faecalis; TE3581; FERM P14237; creatineamidinohydrolase; ds.

XX Alkaligenes faecalis.

XX Key location/qualifiers

FT CDS 1..1212

FT /\*tag= a

FT /transl\_except= (pos:433..435,aa:Glu)

FT /product= "creatineamidinohydrolase"

FT /partial

FT /note= "CDS lacks a stop codon"

XX JP2001252088-A.

XX 18-SEP-2001.

XX 16-MAY-1995; 2001JP-00051054.

XX 16-MAY-1995; 95JP-00117283.

XX (TOYM ) TOYOBOK K.

XX WPI; 2002-003140/01.

XX P-PSDB; AAM51471.

XX A gene encoding creatineamidinohydrolase.

XX Claim 4; Page 10; 11pp; Japanese.

XX The invention relates to Alkaligenes faecalis TE3581 (FERM P-14237)

XX CC creatineamidinohydrolase and the encoding gene. The gene can be used for

XX CC the commercial preparation of creatineamidinohydrolase. (Updated on 07-

XX AUG-2003 to correct OS field.)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

XX Query Match 99.9%; Score 1210.4; DB 6; Length 1212;

XX Best Local Similarity 99.9%; Pred. No. 1.5e-195;

XX Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 ATGATGACGACATGTTGACGATGTAATGACAAACGGGAGAAAATTATTCGCCG 60

```

```

Db      1 ATGACTGACGACATGTTGCACTGTGTAATGSCAACAGCGCGAGAAAGTATTTCCCG 60
Qy      61 TTTTTCGGATGCGGATGATGACCCCGCCCAAAAGAGCTTGGCGCTGGATGGCCAGAAC 120
Db      61 TTTTTCGGATGCGGATGATGACCCCGCCCAAAAGAGCTTGGCGCTGGATGGCCAGAAC 120
Qy      121 AATGTCGATGCGGCGCTGTTCACTCTTTATCACTGCATCAACTACTATTCGCGCTGGCTG 180
Db      121 AATGTCGATGCGGCGCTGTTCACTCTTTATCACTGCATCAACTACTATTCGCGCTGGCTG 180
Qy      181 TACTGCTATTTGCGACGCAAGTACGCGCATGATCGACCAACAACGCGCAGACGATT 240
Db      181 TACTGCTATTTGCGACGCAAGTACGCGCATGATCGACCAACAACGCGCAGACGATT 240
Qy      241 TCGGCGCGGATCGACGCGCGCGCAGCCCTGGCGCGCTTCGCGCGCAACATCACTTAC 300
Db      241 TCGGCGCGGATCGACGCGCGCGCAGCCCTGGCGCGCTTCGCGCGCAACATCACTTAC 300
Qy      301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGGCGCGCAGCTGACCAAGCGCGCAG 360
Db      301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGGCGCGCAGCTGACCAAGCGCGCAG 360
Qy      361 CGCATCGGCGATCGAGTTGCAACAAGTCAATCTGACTTCCGCGCGCAGCTCGAGAGGCC 420
Db      361 CGCATCGGCGATCGAGTTGCAACAAGTCAATCTGACTTCCGCGCGCAGCTCGAGAGGCC 420
Qy      421 CTACCGGCGCTCGACTTCTGTCGACATGACGCAACCTCGATGATGATGCGCAGCATCAAG 480
Db      421 CTACCGGCGCTCGACTTCTGTCGACATGACGCAACCTCGATGATGATGCGCAGCATCAAG 480
Qy      481 TCGCTGGAAGAGAGAAAGCTGATCCGGAAGGCGCGCGCTGATGATGATGCGCGCGCG 540
Db      481 TCGCTGGAAGAGAGAAAGCTGATCCGGAAGGCGCGCGCTGATGATGATGCGCGCGCG 540
Qy      541 GCCTGCGCGGCTGCTCAACAAGCGCGCGCTGCGAGATGAAGTGGCGATCGCCACACC 600
Db      541 GCCTGCGCGGCTGCTCAACAAGCGCGCGCTGCGAGATGAAGTGGCGATCGCCACACC 600
Qy      601 AATGCGATGATCGCGGATGATCGCAAAATGCTTCCCTTCTGGAAGTATGAGACACTGG 660
Db      601 AATGCGATGATCGCGGATGATCGCAAAATGCTTCCCTTCTGGAAGTATGAGACACTGG 660
Qy      661 ACCTGCTTCAGTCCGCGCATCAACACGCGCGCGCAATCGCGGATGATGATGATGATGATG 720
Db      661 ACCTGCTTCAGTCCGCGCATCAACACGCGCGCGCAATCGCGGATGATGATGATGATGATG 720
Qy      721 GTGCAATCGCGGACATCTTTTCTGCTCAACACTTCCGATGATCTTCCGCTACTACCC 780
Db      721 GTGCAATCGCGGACATCTTTTCTGCTCAACACTTCCGATGATCTTCCGCTACTACCC 780
Qy      781 GCGCTGGAAGCGCGCTGTTCTGCGACATGATGATGATGATGATGATGATGATGATGATG 840
Db      781 GCGCTGGAAGCGCGCTGTTCTGCGACATGATGATGATGATGATGATGATGATGATGATG 840
Qy      841 AAGAAGTGCGCGCTGATCGCGCGCGCGCTGAGCTATCAAGCGCGCGCGCGCTGCAAG 900
Db      841 AAGAAGTGCGCGCTGATCGCGCGCGCGCTGAGCTATCAAGCGCGCGCGCGCTGCAAG 900
Qy      901 GACATGCGCATGAGCTCAAGAGATGTAACGAGATGGAAGCTTCTGTAAGTACCGCTCC 960
Db      901 GACATGCGCATGAGCTCAAGAGATGTAACGAGATGGAAGCTTCTGTAAGTACCGCTCC 960
Qy      961 TTGCGGTTATGCGCACTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db      961 TTGCGGTTATGCGCACTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy      1021 GAGCTGCGGAGGACATCGACACGAGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db      1021 GAGCTGCGGAGGACATCGACACGAGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy      1081 ATGCTGATGCTGCGGAGGCGATGCGCGGCTGCGCGGCTGATGCGAGCAGCATCTTG 1140
Db      1081 ATGCTGATGCTGCGGAGGCGATGCGCGGCTGCGCGGCTGATGCGAGCAGCATCTTG 1140

```

```

Qy      1141 ATGCTGCGGAGGACGCTGCCGAGAACATCACCGGCTTCCGTTCCGTCGCGACACACC 1200
Db      1141 ATGCTGCGGAGGACGCTGCCGAGAACATCACCGGCTTCCGTTCCGTCGCGACACACC 1200
Qy      1201 ATCATCGCGCAC 1212
Db      1201 ATCATCGCGCAC 1212

RESULT 6
ABA93696
ID ABA93696 standard; DNA; 1212 BP.
XX
AC ABA93696;
XX
XX 29-APR-2002 (first entry)
XX
DE Creatine amidinohydrazase encoding DNA SEQ ID NO:2.
XX
XX Creatine amidinohydrazase; enzyme; mutant; stable; clinical diagnosis;
XX gene; da.
XX
OS Alcaligenes faecalis.
XX
FH Key Location/Qualifiers
FT CDS 1..1212
FT /*tag= a
FT /product= "creatine amidinohydrazase"
FT /transl_except= (pos:433..435,aa:Glu)
FT /note= "no stop codon given"
FT
PN JP2001346594-A.
XX
XX 18-DEC-2001.
XX
XX 17-DEC-1996; 2001JP-00121708.
XX
XX 17-DEC-1996; 96JP-00337027.
XX
XX (TOYOBO KK.
XX
DR MPI: 2002-145187/19.
XX P-PSDB; ABB05660.
XX
XX Mutant creatine amidinohydrazase for use in clinical diagnosis has long-
XX term stability in a neutral buffer solution.
XX
PS Claim 12; Page 12-13; 15pp; Japanese.
XX
XX
XX The present invention describes a stable mutant creatine amidinohydrazase
XX having a long-term stability in a neutral buffer solution. Also described
XX are: (1) a creatine amidinohydrazase gene encoding the above stable
XX creatine amidinohydrazase; (2) a gene encoding a mutant creatine
XX amidinohydrazase having a long-term stability in a neutral buffer solution
XX compared to wild type creatine amidinohydrazase, in which at least one gene
XX among those found in a fully defined 1212 nucleotide sequence (the
XX present sequence) is replaced by another gene; (3) a recombinant plasmid
XX containing a gene encoding the above creatine amidinohydrazase; (4) a cell
XX transformed by the above plasmid; and (5) a method for the preparation of
XX a stable creatine amidinohydrazase in which the above cell is cultured in a
XX medium and creatine amidinohydrazase is collected. The creatine
XX amidinohydrazase is useful as a clinical diagnosing agent
XX
SQ Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 1.5e-195;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGACTGACGACATGTTGCACTGTGTAATGSCAACAGCGCGAGAAAGTATTTCCCG 60
Db      1 ATGACTGACGACATGTTGCACTGTGTAATGSCAACAGCGCGAGAAAGTATTTCCCG 60

```



```

QY 61 TTTTGGATGCGGAGATGACCCGCGCGCAAAACGAGCTTCGCGCTGGAATGGCCAGAAC 120
DB 61 TTTTGGATGCGGAGATGACCCGCGCGCAAAACGAGCTTCGCGCTGGAATGGCCAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCAACTATTCGGGCTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCAACTATTCGGGCTGGCTG 180
QY 181 TACTGCTATTTGGAGCGAAGTACGCGATGTCATGACACAAACACCGCAGAGATT 240
DB 181 TACTGCTATTTGGAGCGAAGTACGCGATGTCATGACACAAACACCGCAGAGATT 240
QY 241 TGGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TGGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TCGCTCGAAGAGCGAAGTGAATCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TCGCTCGAAGAGCGAAGTGAATCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 601 AATGCGATGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 661 ACTGCGTTCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACTGCGTTCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 GTGCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 GCGCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 GCGCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 AAGAACTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAACTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 901 GACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 961 TTGCGGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 TTGCGGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1021 GAGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 GAGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 AATGATGATGCTGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1081 AATGATGATGCTGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140

```

```

QY 1141 ATCGTCGGGAGGAGCGGTCGCGAGAACATCACCGGCTTCGCGTTCGCGGACACACC 1200
DB 1141 ATCGTCGGGAGGAGCGGTCGCGAGAACATCACCGGCTTCGCGTTCGCGGACACACC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 7
AAT13291
ID AAT13291 standard; DNA; 1215 BP.
XX
AC AAT13291;
XX
DT 16-OCT-2003 (revised)
DT 26-SEP-1996 (first entry)
XX
DE Creatine amidinohydrolyase.
XX
KW Creatine amidinohydrolyase; CAH; kidney; disease; Alcaligenes; de.
XX
OS Alcaligenes sp.; KS-85 PERM BP-4487.
XX
FH Key Location/Qualifiers
FT misc_difference 319..321
FT /tag=a
FT /transl_except= 319..321; aa:asn
XX
PN DE19536506-A1.
XX
PD 04-APR-1996.
XX
PF 29-SEP-1995; 95DE-01036506.
XX
PR 29-SEP-1994; 94JP-00235737.
XX
PA (KIKK ) KIKKOMAN CORP.
XX
PI Furukawa K, Ichikawa T, Suzuki M, Koyama Y;
XX
DR WPI: 1996-180805/19.
XX
DR P-PSDB; AAR94463.
XX
PT DNA encoding creatine amidinohydrolyase - useful for quantification of
PT creatine to, e.g. diagnose kidney disease.
XX
PS Claim 1; Page 9-11; 18pp; German.
XX
CC CAH is used for quantification of creatine, e.g. to diagnose kidney
CC disease by measuring creatine content of serum or urine. CAH DNA can be
CC inserted into host cells for the prodn. of CAH. CAH can now be produced
CC efficiently without having to add creatine to the culture medium.
CC (updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 1215 BP; 247 A; 398 C; 357 G; 213 T; 0 U; 0 Other;

Query Match 97.8%; Score 1184.8; DB 2; Length 1215;
Best Local Similarity 98.6%; Pred. No. 3-2e-191;
Matches 1195; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATTTTCACGTGATGAAATGCGCAACGCGGAGAAAGATTATTCGCCG 60
DB 1 ATGACTGACGACATTTTCACGTGATGAAATGCGCAACGCGGAGAAAGATTATTCGCCG 60
QY 61 TTTTGGATGCGGAGATGACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 TTTTGGATGCGGAGATGACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCAACTATTCGGGCTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCAACTATTCGGGCTGGCTG 180

```

```

QY 181 TACGCTATTTCGAGCGCAGTACGGCATGTCATCGACCAACAACCGCCAGCGATT 240
DB 181 TACTGCTATTTCGAGCGCAGTACGGCATGTCATCGACCAACAACCGCCAGCGATT 240
QY 241 TCGGCGCGGATCGACCGCGCGCAGCCCTGGCGCGCAGCTTCGCGCAACAATCACTAC 300
DB 241 TCGGCGCGGATCGACCGCGCGCAGCCCTGGCGCGCAGCTTCGCGCAACAATCACTAC 300
QY 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGCTGGCGCACTGACACAGCGCGCAAG 360
DB 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGCTGGCGCACTGACACAGCGCGCAAG 360
QY 361 CGCATCGGATCGAGTTCGACCACTGCACTTCCGCGCGCACTGCGCGCAAGAAAC 420
DB 361 CGCATCGGATCGAGTTCGACCACTGCACTTCCGCGCGCACTGCGCGCAAGAAAC 420
QY 421 CTACCGGCGCTGACTTCGTGCAATCAGCCAGCCCTCGATGATGATGCGCACATCAAG 480
DB 421 CTACCGGCGCTGAGTTCGTGCAATCAGCCAGCCCTCGATGATGATGCGCACATCAAG 480
QY 481 TCGCTCGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGTGTGATAGCTCGCGCGCG 540
DB 481 TCGCTCGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGTGTGATAGCTCGCGCGCG 540
QY 541 GCCTGCGCGGCTGCTCAACAGCGCGGCTGCCGAGCATGAATGGCGATCGCCACACC 600
DB 541 GCCTGCGCGGCTGCTCAACAGCGCGGCTGCCGAGCATGAATGGCGATCGCCACACC 600
QY 601 AATGCGATGATCCGCGAGATCGCCAAATCGTTCCCTTCGTGAGCTGATGAGACACTGG 660
DB 601 AATGCGATGATCCGCGAGATCGCCAAATCGTTCCCTTCGTGAGCTGATGAGACACTGG 660
QY 661 ACCTGTTTCCAGTCGGGCTCAACAACCGAGCGCGCAACAATCCGCTCACCAACCGCATC 720
DB 661 ACCTGTTTCCAGTCGGGCTCAACAACCGAGCGCGCAACAATCCGCTCACCAACCGCATC 720
QY 721 GTGCAATCCGCGGACATCTTTCGCTCAACACCTTCCCGATGATCTTCCGCTACTACCC 780
DB 721 GTGCAATCCGCGGACATCTTTCGCTCAACACCTTCCCGATGATCTTCCGCTACTACCC 780
QY 781 GCGCTGGAGCGCAGCTGTTCTGCGACATGTCATGACCGCCAGCTCGACATCTGGAG 840
DB 781 GCGCTGGAGCGCAGCTGTTCTGCGACATGTCATGACCGCCAGCTCGACATCTGGAG 840
QY 841 AAGAATCGGCGCTGCTGATCGCCGCGGCTGAGCTGATCAACCGCGCGCGCTCGAAG 900
DB 841 AAGAATCGGCGCTGCTGATCGCCGCGGCTGAGCTGATCAACCGCGCGCGCTCGAAG 900
QY 901 GACATCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTGCTGAAGTACCGGCTCC 960
DB 901 GACATCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTGCTGAAGTACCGGCTCC 960
QY 961 TTGCGCTATGCGCACTCTTTCGCGCTGCTGTGCTACTACTACGCTGCGAGCGCGCTG 1020
DB 961 TTGCGCTATGCGCACTCTTTCGCGCTGCTGTGCTACTACTACGCTGCGAGCGCGCTG 1020
QY 1021 GAGCTCGGAGGAGCATCGACACCGAGCTGAAGCCCGGAGTGTGCTTCATGAGCGCG 1080
DB 1021 GAGCTCGGAGGAGCATCGACACCGAGCTGAAGCCCGGAGTGTGCTTCATGAGAGCGG 1080
QY 1081 ATGGTATGCTGCGAGGAGCATGCCGAGTGGCGGCGGCTATGCGAGCAGCATCTCG 1140
DB 1081 ATGGTATGCTGCGAGGAGCATGCCGAGTGGCGGCGGCTATGCGAGCAGCATCTCG 1140
QY 1141 ATGCTGCGGAGGAGCGGTGCGAGAAATCAACCGGCTTCCGCTTCGCTCGAGAACACC 1200
DB 1141 ATGCTGCGGAGGAGCGGTGCGAGAAATCAACCGGCTTCCGCTTCGCTCGAGAACACC 1200
QY 1201 ATCATTCGCAAC 1212
DB 1201 ATCATTCGCAAC 1212

```

```

RESULT 8
ACC69514
ID ACC69514 standard; DNA; 1215 BP.
XX
AC ACC69514;
XX
DT 21-JUL-2003 (first entry)
XX
DE Erwinia sp. (DSM 97-934) creatinase encoding DNA SEQ ID NO:1.
XX
KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KW chronic nephritis; acute nephritis; tonic muscular dystrophy; gene; ds.
XX
OS Erwinia sp.
XX
FH Key Location/Qualifiers
FT CDS 1..1215
FT /tag= a
FT /BC_number= "3.5.3.3"
FT /product= "creatinase"
XX
PN EP1298213-A1.
XX
PD 02-APR-2003.
XX
PF 17-SEP-2002; 2002EP-00020793.
XX
PR 20-SEP-2001; 2001EP-00121780.
XX
PA (HOFF) ROCHE DIAGNOSTICS GMBH.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
PI
XX WPI; 2003-383834/37.
XX DR P-PSDB; ABR43467.
XX
PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinine and/or creatine concentration in a sample.
XX
PS Disclosure; Page 15-17; 51pp; English.
XX
CC The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I304 and F395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and/or creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine: they are much better
CC suited to detection methods for creatine. The present sequence encodes
CC Erwinia sp. (DSM 97-934) wild-type creatinase from the present invention
XX
SQ Sequence 1215 BP; 242 A; 392 C; 358 G; 223 T; 0 U; 0 Other;
XX
Query Match 81.0%; Score 981.6; DB 8; Length 1215;
Best Local Similarity 88.1%; Pred. No. 6; 5e-157;
Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 1 ATGACTGACGACATTTGACGTGATGAATGGCAACGCGGAGAGAAATATTTCGCCG 60
DB 1 ATGACTGACGACATTTGACGTGATGAATGGCAACATGTGTGAAGAAATATTCGCCG 60
QY 61 TTTTCGATGCGAGATGACCGCGCGCAACGAGCTTGGCGGTGATGCGCAAGAAC 120
DB 61 TTTTCGATGCGAGATGACCGCGCGCAAGTGAAGTGGCGGCTGATGCGCGCAAGAAC 120

```

QY 121 AATGCGATGCGGCGCTGTTCACTTCTATCACTGCACTAATTCGCGCTG 180  
 DB 121 GAGCTGAGCGCTGCTGTTCACTTCTATCACTGCACTAATTCGCGCTG 180  
 QY 181 TACTGCTATTTGCGAGCGAAGTACGCGATGTCATGACCAACAACGCGACGAGATT 240  
 DB 181 TACTGCTATTTGCGCGCAATAACGCGATGTCATGACCAACAACGCGACGAGATT 240  
 QY 241 TTGCGCGGCGATGACGCGCGCGCAAGCCCTGCGCGCGCGAGCTTTCGCGCAACATCACTAC 300  
 DB 241 TTGCGCGGCGATGACGCGCGCGCGCAAGCCCTGCGCGCGCGAGCTTTCGCGCAACATCACTAC 300  
 QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCTGCGCGCGCTGACCAACGCGCGCGCGAG 360  
 DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCTGCGCGCGCTGACCAACGCGCGCGCGAG 360  
 QY 361 CGCATCGCGATGAGTTCGACCAAGTCAATCTGCACTTCGCGCGCGAGCTCGAGGAAAGCC 420  
 DB 361 CGCATCGCGATGAGTTCGACCAAGTCAATCTGCACTTCGCGCGCGAGCTCGAGGAAAGCC 420  
 QY 421 CTACCGGCGCTGACTTGTGCAATCAAGCCGCTGATGTGATGCGACATCAAG 480  
 DB 421 CTGCGCGGCGTGTGATGCTGATATCGGTCAACGCTGATGTGATGCGACAGGTCAG 480  
 QY 481 TTGCGTGAAGAGCAAGCTGATCCGCGAAGGCGCGCGTGTGTGAGTGTGCGCGCGCGCG 540  
 DB 481 TTGCGTGAAGAGCAAGCTGATCCGCGAAGGCGCGCGTGTGTGAGTGTGCGCGCGCGCG 540  
 QY 541 GCCTGCGCGCGCTGCCATCAAGCGCGCGTCCGAGCAATGAGTGTGCGCGCGCGCG 600  
 DB 541 GCCTGCGCGCGCTGCCATCAAGCGCGCGTCCGAGCAAGGTCGCGATGCGCGCGCGCG 600  
 QY 601 AATGCGATGATCCGCGAATGCGCAATTCCTTCTGTGAGCTGATGAGCACTTGG 660  
 DB 601 AATGCGATGATCCGCGAATGCGCAATTCCTTCTGTGAGCTGATGAGCACTTGG 660  
 QY 661 AACTGCTTCCAGTCCGCGATCAACAACGCGCGCGCGCAATTCGCGCGCGCGCG 720  
 DB 661 AACTGCTTCCAGTCCGCGATCAACAACGCGCGCGCGCAATTCGCGCGCGCGCG 720  
 QY 721 GTGCAATCCGCGCATCTTTCGCTCAACACTTCGCGATGATCTTTCGCGCTACTAC 780  
 DB 721 GTGCAATCCGCGCATCTTTCGCTCAACACTTCGCGATGATCTTTCGCGCTACTAC 780  
 QY 781 GCGCTGAGCGCAGCGCTTTCGCGACATGTGATGACGCGCGCTGCACTTGGGAG 840  
 DB 781 GCGCTGAGCGCAGCGCTTTCGCGACATGTGATGACGCGCGCTGCACTTGGGAG 840  
 QY 841 AAGAACGTGCGCGCTGATCGCGCGCGCTCGAGCTGATCAACGCGCGCGCGCGCTG 900  
 DB 841 AAGAACGTGCGCGCTGATCGCGCGCGCTCGAGCTGATCAACGCGCGCGCGCGCTG 900  
 QY 901 GACATCGCGATGAGCTCAAGAGATGTAACGCGCGAGTGGAGACTTGAATACGCGCTC 960  
 DB 901 GATATCGCGATGAGCTCAAGAGATGTAACGCGCGAGTGGAGACTTGAATACGCGCTC 960  
 QY 961 TTGCGCTATGCGCACTCTTCGCGCTGCTGTGCGCACTAATCGATGCGCGAGCGCGCTG 1020  
 DB 961 TTGCGCTATGCGCACTCTTCGCGCTGCTGTGCGCACTAATCGATGCGCGAGCGCGCTG 1020  
 QY 1021 GAGCTGCGCGAGAGATGACACGCGAGTGAAGCCCGCGATGTGTCTTCATGAGCGCG 1080  
 DB 1021 GAACTGCGCGAGAGATGACACGCGAGTGAAGCCCGCGATGTGTCTTCATGAGCGCG 1080  
 QY 1081 ATGATGATGTTGCGCGAGAGGCGCTCCGCGCGCGCGCTACCGCGACAGACATCTCTG 1140  
 DB 1081 ATGATGATGTTGCGCGAGAGGCGCTCCGCGCGCGCGCTACCGCGACAGACATCTCTG 1140  
 QY 1141 ATGCTGCGGAGAGAGCGTGTGCGAGAAATCAACCGGCTTCCGATGTGCTCGGAAACACAC 1200  
 DB 1141 ATGCTGCGGAGAGAGCGTGTGCGAGAAATCAACCGGCTTCCGATGTGCTCGGAAACAC 1200

QY 1201 ATCATCCGCAAC 1212  
 DB 1201 ATCATCCGCAAC 1212  
 RESULT 9  
 ACC69519  
 ID ACC69519 standard; DNA, 1215 BP.  
 XX  
 AC ACC69519;  
 XX  
 DT 21-JUL-2003 (first entry)  
 XX  
 DE Mutant Erwinia creatinase CTGc2 encoding DNA SEQ ID NO:17.  
 XX  
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;  
 KW Erwinia-type creatinase; creatine; creatinase; ureamase; gigantism;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;  
 XX gene; ds.  
 XX  
 OS Erwinia sp.  
 OS Synthetic.  
 XX  
 FH Key  
 FT CDS  
 FT  
 FT  
 FT  
 FT  
 PN EP1298213-A1.  
 XX  
 PD 02-APR-2003.  
 XX  
 PF 17-SEP-2002; 2002EP-00020793.  
 XX  
 PR 20-SEP-2001; 2001EP-00121780.  
 XX  
 PA (HOFF ) ROCHE DIAGNOSTICS GMBH.  
 PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX  
 PI Shao Z, Schmuck R, Kratzsch P, Kenkles J, Weisser H;  
 XX WPI: 2003-383834/37.  
 DR P-PSDB; ABR43478.  
 XX  
 PT New variant of an Erwinia-type creatinase modified relative to a wild-  
 PT type creatinase having creatinase activity, useful for determining  
 PT creatinase and/or creatine concentration in a sample.  
 XX  
 XX Example 4; Page 33-35; 51pp; English.  
 PS  
 CC The present invention describes a variant of an Erwinia-type creatinase  
 CC (I) modified relative to a wild-type creatinase having creatinase  
 CC activity. The variant comprises at least one amino acid substitution at a  
 CC position of the fully defined 404 amino acid sequence given in ABR43467;  
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has  
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.  
 CC Also described is a reagent (II) for determining creatine comprising the  
 CC Erwinia-type creatinase variant. The variant is useful for determining  
 CC creatinine and/or creatine concentration in a sample. Measuring  
 CC creatinine and creatine are useful for diagnosing uremia, chronic  
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some  
 CC other related diseases. The mutant enzymes have improved stability, lower  
 CC conductivity and/or lower Km-values for creatine; they are much better  
 CC suited to detection methods for creatine. The present sequence encodes a  
 CC mutant Erwinia creatinase from the present invention  
 XX  
 SO Sequence 1215 BP; 241 A; 392 C; 361 G; 221 T; 0 U; 0 Other;  
 Query Match 81.0%; Score 981.6; DB 8; Length 1215;  
 Best Local Similarity 88.1%; Pred. No. 6; 5e-157;  
 Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTTGCATGATGAAATGGCAACGCGGAGAAATTTATTCGCCG 60  
 Db 1 ATATGACGACATGTTTGCATGATGAAATGGCAATGATGAGAAATTTATTCGCCG 60  
 QY 61 TTTTGGATGCGGATGACCCGCGCAAAAAGACATTCGGGCTGATGCGCAAGAAC 120  
 Db 61 TTTTCCGATCCGATGACGCGCGCGCAAGTGAAGTGGGCGCTGATGCGCAAGAAC 120  
 QY 121 AATGTCGATGCGGCGCTGTTACCTTATCACTGATCAATCTATTCGGGCTGCTG 180  
 Db 121 GAGTCGAGCGCTGCGCTGTTACCTTATCACTGATCAATCTATTCGATGATGCTG 180  
 QY 181 TACTGCTATTTGGAGCGAAGTACGAGATGATGACCAACAACGCGCAAGATTT 240  
 Db 181 TACTGCTATTTGGAGCGAAGTACGAGATGATGACCAACAACGCGCAAGATTT 240  
 QY 241 TCGGCGGCGCATCGACGCGCGCGAGCCCTGCGCGCGGAGCTTCGCGCAACATCATCTAC 300  
 Db 241 TCGGCGGCGCATCGATGCGCGCTAGCCCTGCGCGCGGAGCTTCGCGCAACATCATCTAT 300  
 QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGCGCAAGCTGACCAAGCGCGCAAG 360  
 Db 301 ACCGACTGCGCGCGCGCAACCTTCTACCAAGCGCGCTGCGCGCAACCTGCGCGCGCAAG 360  
 QY 361 CGGATGCGCATCGAGTTTCCACACGTCATCTGATTCGCGCGCGCAAGCTTCGAGAACGC 420  
 Db 361 CGGATGCGCATCGAGTTTCCACACGTCATCTGATTCGCGCGCGCAAGCTTCGAGAACGC 420  
 QY 421 CTACCGGCGCGCTGACTTCTGTCGATCAGCGACGCGCTGATGTCGATGTCGACCATCAAG 480  
 Db 421 CTACCGGCGCGCTGACTTCTGTCGATCAGCGACGCGCTGATGTCGATGTCGACCATCAAG 480  
 QY 481 TCGCTCGAAGACGAGATGATTCGCGAAGCGCGCGCTGTCGATCGCGCGCGCG 540  
 Db 481 TCGCTCGAAGACGAGATGATTCGCGAAGCGCGCGCTGTCGATCGCGCGCGCG 540  
 QY 541 GCGTGGCGCGCTGCTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCTGTCGATCGCGCGCG 600  
 Db 541 GCGTGGCGCGCTGCTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCTGTCGATCGCGCGCG 600  
 QY 601 AATGCGATGATTCGCGGAGATTCGCAATCGTTCCCTTCTGTCGATGATGATGACACTCTG 660  
 Db 601 AATGCGATGATTCGCGGAGATTCGCAATCGTTCCCTTCTGTCGATGATGATGACACTCTG 660  
 QY 661 ACCTGTTCAGTTCGCGGATCAACACGACGCGCGCGCAACATCGGTCACCAACGCGATC 720  
 Db 661 ACCTGTTCAGTTCGCGGATCAACACGACGCGCGCGCAACATCGGTCACCAACGCGATC 720  
 QY 721 GTGCGATTCGCGGACATCTTCTTGGCTCAACCTTCCGATGATCTTCCGCTACTACAC 780  
 Db 721 GTGCGATTCGCGGACATCTTCTTGGCTCAACCTTCCGATGATCTTCCGCTACTACAC 780  
 QY 781 GCGCTGGAGCGACGCTGTTCTGACACATGTCGATGACGCGCGCTGCAACATCTGAGGAG 840  
 Db 781 GCGCTGGAGCGACGCTGTTCTGACACATGTCGATGACGCGCGCTGCAACATCTGAGGAG 840  
 QY 841 AAGAACGTGCGCTGTCATGCGCGCGCTGTCGATGATCAAGCGCGCGCGCTGCAAG 900  
 Db 841 AAGAACGTGCGCTGTCATGCGCGCGCTGTCGATGATCAAGCGCGCGCGCTGCAAG 900  
 QY 901 GACATGCGCATGAGTCAACGAGATGTCACGCGGATGGGACTGCTGTAAGTACCGCTCC 960  
 Db 901 GATATGCGCATGAGTCAACGAGATGTCACGCGGATGGGATGCTGTAAGTACCGCTCC 960  
 QY 961 TTGCGGCTATGCGCATCTTCTGCGCGCTGTCGATGATCAAGCGCGCGCGCGCTG 1020  
 Db 961 TTGCGGCTATGCGCATCTTCTGCGCGCTGTCGATGATCAAGCGCGCGCGCGCTG 1020  
 QY 1021 GAGCTGCGGAGACATCGACACGAGCTGAAGCGCGCGGATGCTTCTCATGAGCGG 1080  
 Db 1021 GAACTGCGGAGACATCGATACCGGTGACGCGCGGATGCTTCTCATGAGCGG 1080  
 QY 1081 ATGTGTATGCTGCGGAGGCGATGCGCGGATGCGCGGCGCTATGCGAGACACATCTCTG 1140

Db 1081 ATGTGTATGCTGCGGAGGCGCTGCGCGGCGGCTACCGGAGACACATCTCTG 1140  
 QY 1141 ATGTCGGGAGAGAGCGTCCGAGAACATCACCGCTTCCGTCGTCGCGAACACACC 1200  
 Db 1141 ATGTCGGGAGAGAGCGGCGAGAACATTTACCGATTCCTTGGGCGCTGAGCAAC 1200  
 QY 1201 ATCATCCGCAAC 1212  
 Db 1201 ATCATCCGCAAC 1212  
 RESULT 10  
 ACC69515  
 ID ACC69515 standard; DNA; 1215 BP.  
 XX  
 AC ACC69515;  
 XX  
 DT 21-JUL-2003 (first entry)  
 XX  
 DE Mutant Erwinia creatinase CTIm24 encoding DNA SEQ ID NO:9.  
 XX  
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;  
 XX Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;  
 KW gene; ds.  
 XX  
 OS Erwinia sp.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..1215  
 FT /\*tag= "a"  
 FT /EC number= "3.5.3.3"  
 FT /product= "creatinase mutant CTIm24"  
 XX  
 PN BP1298213-AL.  
 PD 02-APR-2003.  
 XX  
 PD 17-SEP-2002; 2002EP-00020793.  
 PF 20-SEP-2001; 2001EP-00121780.  
 XX  
 PR (HOFF ) ROCHE DIAGNOSTICS GMBH;  
 PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX  
 PI Shao Z, Schmuck R, Kratzsch P, Kenkles J, Weisner H;  
 XX WPI; 2003-383834/37.  
 DR P-PSDB; ABR43474.  
 XX  
 PT New variant of an Erwinia-type creatinase modified relative to a wild-  
 PT type creatinase having creatine activity, useful for determining  
 PT creatinine and/or creatine concentration in a sample.  
 XX  
 PS Example 4; Page 20-21; 51dp; English.  
 XX  
 CC The present invention describes a variant of an Erwinia-type creatinase  
 CC (I) modified relative to a wild-type creatinase having creatinase  
 CC activity. The variant comprises at least one amino acid substitution at a  
 CC position of the fully defined 404 amino acid sequence given in ABR43467;  
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has  
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.  
 CC Also described is a reagent (II) for determining creatine comprising the  
 CC Erwinia-type creatinase variant. The variant is useful for determining  
 CC creatinine and/or creatine concentration in a sample. Measuring  
 CC creatinine and creatine are useful for diagnosing uraemia, chronic  
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some  
 CC other related diseases. The mutant enzymes have improved stability, lower  
 CC conductivity and/or lower Km-values for creatine; they are much better  
 CC suited to detection methods for creatine. The present sequence encodes a  
 CC mutant Erwinia creatinase from the present invention

XX Sequence 1215 BP; 241 A; 393 C; 359 G; 222 T; 0 U; 0 Other;  
 SQ Query Match 80.7%; Score 978.4; DB 8; Length 1215;  
 Beat Local Similarity 88.0%; Pred. No. 2.2e-156;  
 Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 1 ATGACGACGACATGTTGCACTGATGAATGACACAGGCGGAGAAAGTTATTTGGCG 60  
 DB 1 ATGACGACGACATGTTGCACTGATGAATGACACAGGCGGAGAAAGTTATTTGGCG 60  
 QY 61 TTTTCGATGCGCAGATGACCCGCGCACAAGAGTTTCGCGCTGATGAGCAAGAAC 120  
 DB 61 TTTTCGATGCGCAGATGACCCGCGCACAAGAGTTTCGCGCTGATGAGCAAGAAC 120  
 QY 121 AATGTCATGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTG 180  
 DB 121 AATGTCATGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTG 180  
 QY 181 TACTGCTATTTGGAGGCAAGTACGGCATGTCATGACCAACAGCGGCAAGATT 240  
 DB 181 TACTGCTATTTGGAGGCAAGTACGGCATGTCATGACCAACAGCGGCAAGATT 240  
 QY 241 TCGGCGCGGATCGACCGCGCGCAGCCCTGGCGCGGAGCTTTCGCGCAACATCACTAC 300  
 DB 241 TCGGCGCGGATCGACCGCGCGCAGCCCTGGCGCGGAGCTTTCGCGCAACATCACTAC 300  
 QY 301 ACCGACTGGCGCGGCAATTTCTATCGCGCGGCGGCACTGACCAAGCGGCGCAAG 360  
 DB 301 ACCGACTGGCGCGGCAATTTCTATCGCGCGGCGGCACTGACCAAGCGGCGCAAG 360  
 QY 361 CGCATGCGATGAGTTTCCATGACGTCGACCTTCCGCGCACTGCGGCAAGAGCG 420  
 DB 361 CGCATGCGATGAGTTTCCATGACGTCGACCTTCCGCGCACTGCGGCAAGAGCG 420  
 QY 421 CTACCGGCGCTGCACTTCTGTCATCAGCAGCAGCCCTGATGATGATGCGCAACATCAAG 480  
 DB 421 CTACCGGCGCTGCACTTCTGTCATCAGCAGCAGCCCTGATGATGATGCGCAACATCAAG 480  
 QY 481 TCGCTCGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGTGACCTCGCGCGCG 540  
 DB 481 TCGCTCGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGTGACCTCGCGCGCG 540  
 QY 541 GCTCGCGCGCTGCGCATCAAGCGCGCGCGCGCGCATGAGTGGCGCATTCGCGACAC 600  
 DB 541 GCTCGCGCGCTGCGCATCAAGCGCGCGCGCGCGCATGAGTGGCGCATTCGCGACAC 600  
 QY 601 AATGCGATGATCGCGAGATCGGCAATGCTTCCCTGCTGAGGCTGATGACACTG 660  
 DB 601 AATGCGATGATCGCGAGATCGGCAATGCTTCCCTGCTGAGGCTGATGACACTG 660  
 QY 661 ACTGCTTCAGTCCGCGATCAACACGAGCGGCGCAACATCCGCTGACCAACCGCATC 720  
 DB 661 ACTGCTTCAGTCCGCGATCAACACGAGCGGCGCAACATCCGCTGACCAACCGCATC 720  
 QY 721 GTGCAATCCGCGACATCTTTTGGCTCAACACTTCCGATGATCTTTCGCTACTACAC 780  
 DB 721 GTGCAATCCGCGACATCTTTTGGCTCAACACTTCCGATGATCTTTCGCTACTACAC 780  
 QY 781 GGGCTGAGAGCGACCGCTGTTTTCGACATGTCATGAGCGCGCTCGACATCTGGAG 840  
 DB 781 GGGCTGAGAGCGACCGCTGTTTTCGACATGTCATGAGCGCGCTCGACATCTGGAG 840  
 QY 841 AAGAACGTCGCGCTGATCGCGCGGCTCGACTGATCAAGCGGCGCGCGCTGCAAG 900  
 DB 841 AAGAACGTCGCGCTGATCGCGCGGCTCGACTGATCAAGCGGCGCGCGCTGCAAG 900  
 QY 901 GACATGCGCATGAGCTCAACGAGATGTCACCGGATGGAATCTGCTGAAGTACCGCTCC 960  
 DB 901 GACATGCGCATGAGCTCAACGAGATGTCACCGGATGGAATCTGCTGAAGTACCGCTCC 960  
 QY 961 TTCGGCTATGCGCACTCTTCGCGCTGCTGTCATCACTACGCGTGCAGGCGCGCGT 1020  
 DB 961 TTCGGCTATGCGCACTCTTCGCGCTGCTGTCATCACTACGCGTGCAGGCGCGCGT 1020

DB 961 TTCGGCTATGCGCACTCTTCGCGCTGCTGTCATCACTACGCGGCGCGCGT 1020  
 QY 1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCCCGGATGATGATCTTCCATGAGCGG 1080  
 DB 1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCCCGGATGATGATCTTCCATGAGCGG 1080  
 QY 1081 ATGTCGATGTCGTCGCGGAGGCGATGCGCGGCTGCGCGCTATGCGAGCAGCATCTG 1140  
 DB 1081 ATGTCGATGTCGTCGCGGAGGCGATGCGCGGCTGCGCGCTATGCGAGCAGCATCTG 1140  
 QY 1141 ATGTCGCGGAGAGCGGTCGCGGAGACATCACTACCGCTTCCGCTGCGGAAACACACC 1200  
 DB 1141 ATGTCGCGGAGAGCGGTCGCGGAGACATCACTACCGCTTCCGCTGCGGAAACACACC 1200  
 QY 1201 ATCATCCGCAAC 1212  
 DB 1201 ATCATCCGCAAC 1212

RESULT 11  
 ACC69517  
 ID ACC69517 standard; DNA; 1212 BP.  
 XX  
 AC ACC69517;  
 XX  
 DT 21-JUL-2003 (first entry)  
 XX  
 DE Mutant Erwinia creatinase CT2m10 encoding DNA SEQ ID NO:13.  
 XX  
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;  
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;  
 KW gene; de.  
 OS Erwinia sp.  
 OS Synthetic.  
 OS  
 FT Key Location/Qualifiers  
 FT CDS 1..1212  
 FT /\*tag= a  
 FT /partial  
 FT /EC number= "3.5.3.3"  
 FT /product= "creatinase mutant CT2m10"  
 FT /note= "no stop codon given"  
 FT  
 PN EP1298213-A1.  
 XX  
 PD 02-APR-2003.  
 XX  
 PF 17-SEP-2002; 2002EP-00020793.  
 XX  
 PR 20-SEP-2001; 2001EP-00121780.  
 XX  
 PA (HOFF ) ROCHE DIAGNOSTICS GMBH.  
 PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX  
 PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisner H;  
 XX  
 DR WPI; 2003-383834/37.  
 DR P-PSDB; ABR43476.  
 XX  
 PT New variant of an Erwinia-type creatinase modified relative to a wild-  
 PT type creatinase having creatinase activity, useful for determining  
 PT creatinine and/or creatine concentration in a sample.  
 XX  
 PS Example 4; Page 26-28; 51pp; English.  
 XX  
 CC The present invention describes a variant of an Erwinia-type creatinase  
 CC (1) modified relative to a wild-type creatinase having creatinase  
 CC activity. The variant comprises at least one amino acid substitution at a  
 CC position of the fully defined 404 amino acid sequence given in ABR43467;  
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has  
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.

CC Also described is a reagent (II) for determining creatine comprising the  
 CC Erwinia-type creatinase variant. The variant is useful for determining  
 CC creatinine and/or creatine concentration in a sample. Measuring  
 CC creatinine and creatine are useful for diagnosing uraemia, chronic  
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some  
 CC other related diseases. The mutant enzymes have improved stability, lower  
 CC conductivity and/or lower Km-values for creatine; they are much better  
 CC suited to detection methods for creatine. The present sequence encodes a  
 CC mutant Erwinia creatinase from the present invention  
 CC  
 XX

Sequence 1212 BP; 240 A; 394 C; 358 G; 220 T; 0 U; 0 Other;

Query Match 80.6%; Score 976.8; DB 8; Length 1212;

Best Local Similarity 87.9%; Pred. No. 4.2e-156;

Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCAAGTAAATGACACACGCGAGAAATATTCGCCG 60  
 DB 1 ATGACTGACGACATGTTGCAAGTAAATGACACACGCGAGAAATATTCGCCG 60  
 QY 61 TTTTCGATCCGAGATGACCCGCCCAAAAGACCTTGGCGCTGGATGCGCAAGAC 120  
 DB 61 TTTTCGATCCGAGATGACCCGCCCAAAAGACCTTGGCGCTGGATGCGCAAGAC 120  
 QY 121 AATGTGATGCGGCGCTTACCTCTTATCAGTCACTCACTACTATTCCGGCTG 180  
 DB 121 AATGTGATGCGGCGCTTACCTCTTATCAGTCACTCACTACTATTCCGGCTG 180  
 QY 181 TACTGCTATTTCCGACGCAAGTACGCGATGATCGACACCAACACGCGACGAT 240  
 DB 181 TACTGCTATTTCCGACGCAAGTACGCGATGATCGACACCAACACGCGACGAT 240  
 QY 241 TCGCGCGGACATGACGCGCGCGACCTTGGCGCGCGAGCTTGGCGCAATCATT 300  
 DB 241 TCGCGCGGACATGACGCGCGCGACCTTGGCGCGCGAGCTTGGCGCAATCATT 300  
 QY 301 ACCGACTGGCGCGCGCAATTTCTATTCGCGCGTGGCGCGACGCGACGCGCG 360  
 DB 301 ACCGACTGGCGCGCGCAATTTCTATTCGCGCGTGGCGCGACGCGACGCGCG 360  
 QY 361 CGGATCGGACATGAGTTGCAACAGTCAATCTGCACTTCCGCGCGACGCTCGA 420  
 DB 361 CGGATCGGACATGAGTTGCAACAGTCAATCTGCACTTCCGCGCGACGCTCGA 420  
 QY 421 CTACCGCGCGCTGCACTTCTGCAATCAGCCAGCTTCAATGAGTGGCAACAT 480  
 DB 421 CTACCGCGCGCTGCACTTCTGCAATCAGCCAGCTTCAATGAGTGGCAACAT 480  
 QY 481 TCGCTGAAAGAGCAGAGCTGATCCGGAAGGCGCGCGTGTGTGACGTGGCGG 540  
 DB 481 TCGCTGAAAGAGCAGAGCTGATCCGGAAGGCGCGCGTGTGTGACGTGGCGG 540  
 QY 541 GCCTGCGGCTGCGCATCAAGGCGCGGTGCGGAGCATGAGTGGCGATCGCCA 600  
 DB 541 GCCTGCGGCTGCGCATCAAGGCGCGGTGCGGAGCATGAGTGGCGATCGCCA 600  
 QY 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGAGACTGATGACAC 660  
 DB 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGAGACTGATGACAC 660  
 QY 661 ACCTGTTCCAGTCGCGCATCAACACGCGCGCGCAATCCGCTGACCAACGCG 720  
 DB 661 ACCTGTTCCAGTCGCGCATCAACACGCGCGCGCAATCCGCTGACCAACGCG 720  
 QY 721 GTGCAATCCGCGCATCTTTTGTCTCAACACCTTCCGATGATCTTCCGCTA 780  
 DB 721 GTGCAATCCGCGCATCTTTTGTCTCAACACCTTCCGATGATCTTCCGCTA 780  
 QY 781 GCGCTGAGAGCGAGCTTCTTCTGCAATGATGAGACGCGCTCGACATCTGGAG 840  
 DB 781 GCGCTGAGAGCGAGCTTCTTCTGCAATGATGAGACGCGCTCGACATCTGGAG 840  
 QY 841 AAGAACGTGGCGGTGATCGCGCGGCTCGAGCTGATCAAGCGCGCGCGCTGCA 900

DB 841 AAGAACGTGGCGGTGATCGCGCGGCTCGAGCTGATCAAGCGCGCGCTGCA 900  
 QY 901 GACATCGCATGAGTCAACAGATGTACCGCGAGTGGAGCTGTGAATACGCTCC 960  
 DB 901 GATATCGCATGAGTCAACAGATGTACCGCGAGTGGAGCTGTGAATACGCTCC 960  
 QY 961 TTGCGTATGAGGCACTCTTCCGCGCTGTGTGCTCACTAAGGTGCGAGCG 1020  
 DB 961 TTGCGTATGAGGCACTCTTCCGCGCTGTGTGCTCACTAAGGTGCGAGCG 1020  
 QY 1021 GAGCTGCGGAGACATCGACACGCGAGTGAAGCCGCGATGATGCTTCCATG 1080  
 DB 1021 GAGCTGCGGAGACATCGACACGCGAGTGAAGCCGCGATGATGCTTCCATG 1080  
 QY 1081 ATGTGATGCTTCCGAGGCGATGCGCGGCTGCGCGGCTATCGCGAGCA 1140  
 DB 1081 ATGTGATGCTTCCGAGGCGATGCGCGGCTGCGCGGCTATCGCGAGCA 1140  
 QY 1141 ATCGTGGGAGAGAGCGTCCGAGAACATCAACGCGCTTCCGCTGCGGAA 1200  
 DB 1141 ATCGTGGGAGAGAGCGTCCGAGAACATCAACGCGCTTCCGCTGCGGAA 1200  
 QY 1201 ATCATCCGCAAC 1212  
 DB 1201 ATCATCCGCAAC 1212

RESULT 12  
 ACC69516  
 ID ACC69516 standard; DNA; 1212 BP.  
 XX  
 AC ACC69516;  
 XX  
 DT 21-JUL-2003 (first entry)  
 XX  
 DE Mutant Erwinia creatinase CT2m9 encoding DNA SEQ ID NO:11.  
 XX  
 KW Erwinia, creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;  
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;  
 KW gene; de.  
 XX  
 OS Erwinia sp.  
 OS Synthetic.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 1..1212  
 FT /\*tag= a  
 FT /partial  
 FT /EC\_number= "3.5.3.3"  
 FT /product= "creatinase mutant CT2m9"  
 FT /note= "no stop codon given"  
 XX  
 PN EP1298213-A1.  
 XX  
 PD 02-APR-2003.  
 XX  
 PF 17-SEP-2002; 2002EP-00020793.  
 XX  
 PR 20-SEP-2001; 2001EP-00121780.  
 XX  
 PA (HOPF ) ROCHE DIAGNOSTICS GMBH.  
 PA (HOPF ) HOPFMANN LA ROCHE & CO AG F.  
 XX  
 PI Shao Z, Schumuck R, Kratzsch P, Kenkles J, Weisner H;  
 DR MPI: 2003-383834/37.  
 DR P-PSDB; ABR43475.  
 XX  
 PT New variant of an Erwinia-type creatinase modified relative to a wild-  
 PT type creatinase having creatinase activity, useful for determining  
 PT creatinine and/or creatine concentration in a sample.



XX Example 4; Page 23-25; s1pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase  
 CC (1) modified relative to a wild-type creatinase having creatinase  
 CC activity. The variant comprises at least one amino acid substitution at a  
 CC position of the fully defined 404 amino acid sequence given in ABR41467;  
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has  
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.  
 CC Also described is a reagent (II) for determining creatine comprising the  
 CC Erwinia-type creatinase variant. The variant is useful for determining  
 CC creatinine and/or creatine concentration in a sample. Measuring  
 CC creatinine and/or creatine are useful for diagnosing uraemia, chronic  
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some  
 CC other related diseases. The mutant enzymes have improved stability, lower  
 CC conductivity and/or lower Km-values for creatine. They are much better  
 CC suited to detection methods for creatine. The present sequence encodes a  
 CC mutant Erwinia creatinase from the present invention

XX Sequence 1212 BP; 239 A; 393 C; 359 G; 221 T; 0 U; 0 Other;

Query Match 80.6%; Score 976.8; DB 8; Length 1212;

Best Local Similarity 87.9%; Pred. No. 4.2e-156;

Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAATGSCACACGCGGAGAAATTTTCGCCG 60  
 DB 1 ATGACTGACGACATGTTGCACTGATGAATGSCACACGCGGAGAAATTTTCGCCG 60  
 QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGCATGTGCGGCGTGTGCGCAAGAC 120  
 DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGCATGTGCGGCGTGTGCGCAAGAC 120  
 QY 121 AATGTGATGCGCGCGTGTTCACCTCTTATCACTGATCACTACTATTCGCGTGC 180  
 DB 121 AATGTGATGCGCGCGTGTTCACCTCTTATCACTGATCACTACTATTCGCGTGC 180  
 QY 122 GAGCTGAGCGTGCCTGCTTCACTCTATCATTTGATCACTACTATTCGCGTGC 180  
 DB 122 GAGCTGAGCGTGCCTGCTTCACTCTATCATTTGATCACTACTATTCGCGTGC 180  
 QY 181 TACTGCTATTTGAGACGCAATGACGATGTCATGACCAACACCCACGACGATT 240  
 DB 181 TACTGCTATTTGAGACGCAATGACGATGTCATGACCAACACCCACGACGATT 240  
 QY 181 TACTGCTATTTGAGACGCAATGACGATGTCATGACCAACACCCACGACGATT 240  
 DB 181 TACTGCTATTTGAGACGCAATGACGATGTCATGACCAACACCCACGACGATT 240  
 QY 241 TGGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
 DB 241 TGGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
 QY 301 ACCGACTGCGCGCGCGCGCAATTTCTATCCGCGCGTGCAGCTGACCAACGCGCG 360  
 DB 301 ACCGACTGCGCGCGCGCGCAATTTCTATCCGCGCGTGCAGCTGACCAACGCGCG 360  
 QY 361 GCGATGCGGATGAGTTCGACCACTGATCTTCCGCGCGCGCGCGCGCGCGCG 420  
 DB 361 GCGATGCGGATGAGTTCGACCACTGATCTTCCGCGCGCGCGCGCGCGCGCG 420  
 QY 421 CTACCGGCGGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 DB 421 CTACCGGCGGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 QY 481 TCGCTGCAAGACGACGATGTCGCGCGAGCGCGCGCGTGTGACGTCGCGCGCG 540  
 DB 481 TCGCTGCAAGACGACGATGTCGCGCGAGCGCGCGCGTGTGACGTCGCGCGCG 540  
 QY 541 GCGTGGCGGCGTGCATGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 DB 541 GCGTGGCGGCGTGCATGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 QY 601 AATGCGATGATGCGCGGAGATGCGCGAAATGCTTCCCTTGTGAACTGATGAC 660  
 DB 601 AATGCGATGATGCGCGGAGATGCGCGAAATGCTTCCCTTGTGAACTGATGAC 660  
 QY 661 ACCTGTTTCAGTTCGCGGATGCAACACGCGCGCGCGCAATTCGCGTGCACCA 720  
 DB 661 ACCTGTTTCAGTTCGCGGATGCAACACGCGCGCGCGCAATTCGCGTGCACCA 720

QY 721 GTGCAATCGGCGGATGATCTTTCGCTCAACCTTCCGAGATGATCTGCGCTACTAC 780  
 DB 721 GTGCAATCGGCGGATGATCTTTCGCTCAACCTTCCGAGATGATCTGCGCTACTAC 780  
 QY 781 GCGCTGAGCGCGACGCTGTTGCGACCATGTCATGACGCGAGCCTGACATCTGGAG 840  
 DB 781 GCGCTGAGCGCGACGCTGTTGCGACCATGTCATGACGCGAGCCTGACATCTGGAG 840  
 QY 841 AAGAACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 DB 841 AAGAACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 QY 901 GACATGCGGATGAGCTGACGAGATGACGCGCGAGTGGAGCTGTAAGATGCGCTCC 960  
 DB 901 GATATGCGGATGAGCTGACGAGATGACGCGCGAGTGGAGCTGTAAGATGCGCTCC 960  
 QY 961 TTTGCGCTATGCGCACTCTTCGCGCGTGTGCTGCGCACTACGCGCGCGCGCG 1020  
 DB 961 TTTGCGCTATGCGCACTCTTCGCGCGTGTGCTGCGCACTACGCGCGCGCGCG 1020  
 QY 1021 GAGCTGCGCGGAGGATGACACCGGAGTGAAGCCCGGACATGCTGCTCCATGAGCC 1080  
 DB 1021 GAGCTGCGCGGAGGATGACACCGGAGTGAAGCCCGGACATGCTGCTCCATGAGCC 1080  
 QY 1081 ATGCTGATGCTGCGCGGAGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
 DB 1081 ATGCTGATGCTGCGCGGAGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
 QY 1141 ATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 DB 1141 ATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 QY 1201 ATCATCGCGCAAC 1212  
 DB 1201 ATCATCGCGCAAC 1212

RESULT 13

ACC69518 standard; DNA; 1215 BP.

ACC69518;

21-JUL-2003 (first entry)

Mutant Erwinia creatinase CT2m28 encoding DNA SEQ ID NO:15.

Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;

Erwinia-type creatinase; creatine; creatinase; uraemia; gigantism;

chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;

gene; ds.

OS Erwinia sp.

XX Synthetic.

XX Key

XX CDS

XX EPI298213-A1.

XX 02-APR-2003.

XX 17-SEP-2002; 2002EP-00020793.

XX 20-SEP-2001; 2001EP-00121780.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX Shao Z, Schmuck R, Kratzsch P, Kenkles J, Weisser H;

Location/Qualifiers  
 1..1215  
 /\*tag= a  
 /EC number= "3.5.3.3"  
 /product= "creatinase mutant CT2m28"

XX WPI; 2003-383834/37.  
DR P-PSDB; ABR43477.  
DR

PT New variant of an Erwinia-type creatinase modified relative to a wild-  
PT type creatinase having creatinase activity, useful for determining  
PT creatinine and/or creatine concentration in a sample.

PS Example 4; Page 30-32; 51pp; English.

The present invention describes a variant of an Erwinia-type creatinase (I) modified relative to a wild-type creatinase having creatinase activity. The variant comprises at least one amino acid substitution at a position of the fully defined 404 amino acid sequence given in ABR43467. There are selected from N130, M203, I278, I330 and F335. Creatinase has the EC number EC 3.5.3 and is also known as creatine amidinohydrolase. CC Also described is a reagent (II) for determining creatine comprising the Erwinia-type creatinase variant. The variant is useful for determining creatinine and/or creatine concentration in a sample. Measuring creatinine and creatine are useful for diagnosing uraemia, chronic nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some other related diseases. The mutant enzymes have improved stability, lower conductivity and/or lower Km-values for creatine: they are much better suited to detection methods for creatine. The present sequence encodes a mutant Erwinia creatinase from the present invention

**SQ** Sequence 1215 BP; 240 A; 394 C; 359 G; 222 T; 0 U; 0 Other;

Query Match	80.6%	Score 976.8	DB 8	Length 1215
Best Local Similarity	87.9%	Pred. NO. 4.2e-156		
Matches 1065, Conservative	0	Mismatches 147	Indels 0	Gaps 0

Qy	1	ATAGCTGACGACATGTTGCGCTGATGAATGGCAACGGGGAAAGATTATTCGCG	60
Db	1	ATAGCTGACGACATGTTGCGCTGATGAATGGCAATGGGTGAAAGAAATTTCCCC	60
Qy	61	TTTTGGATGTCGAGATGACCGCGCGCCAAAAGAGTTGCGGCTGATGGCAAGAC	120
Db	61	TTTTCCGATGCCAGATGACGCGCGCCAGAGTGAAGTGGCGCTGATGGCGAAAC	120
Qy	121	AATGTGATGTCGGCGCTGTTCACTCTTTATCACTGATCAATACTATTCCGCTGGCTG	180
Db	121	GAGTCGACGCTGCGCTGTTCACTCTTCAATGATCAACTACTACTCTGGAATCCTG	180
Qy	181	TACTGCTATTTGCGAGCGCAAGTACGCGCATGTGATCGACACAACAACGCCACGATTT	240
Db	181	TACTGCTATTTGCGGCGCGCAATACGCGCATGTGATCGACAGACATATGCCACATTC	240
Qy	241	TGGCGCGGCATCGACGGCGCGCAGCCCTGGCGCGCGAGCTTGCGGCAACATCACTTAC	300
Db	241	TGGCGCGGCATCGATGGCGGTCAGCCCTGGCGCGGAGCTTGCGGCAACATCACTTAT	300
Qy	301	ACGCACTGGCGCGCGCGCAAAATTTCTATTCGCGCGTGGCGCACTTACCACGGGGCGCAAG	360
Db	301	ACGCACTGGCGCGCGCGCAAACTTTCTATTCAGCGCGCTGGCGCAACTACCCCGGGCGCAGG	360
Qy	361	CGCATCGGCATCGAGTTTCGACACACGTCATTCGACTTCGCGCGCGCAGCTGAGAGAAC	420
Db	361	CGCATCGGTATCGAGTTTCGATCACTGCGTGAACCTTTGACTTCGCGCGCGCAGCTGAGAGAGCG	420
Qy	421	CTACCGGGCGTCGACTTCCTCGACATCAGCAGCCCTGATGTGGATGCGACATCAAG	480
Db	421	CTGCGCGGCGTCGAGTTTCGTCGATATCGGTCAACCGTCAATGTGGATGCGACGAGTCAAG	480
Qy	481	TGCGCTCGAAGAGCAAGAGCTGATTCGCGCGAAGCGCCGCGTGTGTGAACTCGGGGCGCG	540
Db	481	TGCGCTCGAAGAGCAAGAGCTGATTCGCGAGGAGTCCCGTATCTGTGAGAGTGGGCGGTGCC	540
Qy	541	GCGTCGCGCGCTCCATTCAAAGCGCGCGTGGCCCGAGCATGGAATGGCATCGCCACCAAC	600
Db	541	GCGTCGCGTTCGCGCGCTCAAAGCGCGCGTTCGCGAGACAGAGGTGCGATGCGCCACCAAC	600
Qy	601	AATGCGATGATCGCGAGATTCGCAAAATGTTCCCTTCGTGGAGCTGATGAGACCTGG	660

Db	601	AAATCGATGATGCTCCGAGATGCGCAAGTCGTTCCCTTCGTGCACTGATGSAACCTTGG	660
Qy	661	ACCTGGTTCCAGTGGGCAATCAACACGACGGCGCGACAAATCCGCTACCAACCGCATC	720
Db	661	ACCTGGTTCCAGTGGGCAATCAACACGACGGCGCGCAATCCGCTGACCAACCGCATC	720
Qy	721	GTGCATTCGGGCGCAATCCCTTTCCGCTCAACACCTTCCCGATGATCTTCGGCTACTACAC	780
Db	721	GTGCATTCGGGCGGATATCTGTGCTCAACAGTTCCTCCGATGATCTTCGGCTACTACAC	780
Qy	781	GCCTGAGAGCGACAGCTGTTCTGCGACCATGTCGATGACGCCCAAGCTTCGATCTGGAG	840
Db	781	GCCTGAGAGCGAACGCTGTTCTGTGACCAAGTCGACGATGCCAGCTTGTACCTGGAG	840
Qy	841	AAGAACGTGGCCGTGATGCGCCGCGGAGCTGAGCTGATCAAGCCGCGCGCGCTGCAG	900
Db	841	AAGAACGTGCGCGTGACCCGCGCGGCTCGAATCATCAAGCCGCGGTGCGGCTGCAG	900
Qy	901	GACATCCGCATCGAGCTCAACGAGATATACCGCGAGTGGGACCTGCTGAAGTACCGCTCC	960
Db	901	GATATCCGCCCTCGAACTCAACGAGATGTACCGGAGTGGGATCTGCTGAAGTACCGCTCC	960
Qy	961	TTGCGCTATGAGCCACTCTCTTCGGCGTGTGTGCACTACAGGTGCGAGGCCGCGGCTG	1020
Db	961	TTGCGCTACGCGCACTCTCTTCGGGAGTCTCTCCACTACTACGCGCGCGAGGCGGCGTGT	1020
Qy	1021	GAGCTGCGGAGGACATCGACACCGAGCTGAAGCCCGGCACTGCTGATCTTCATGAGACCG	1080
Db	1021	GAACTGGGCGAGGACATCGATACCGTGTCTGACGCGCGGCACTGGTGGCTCTCATGAGACCG	1080
Qy	1081	ATGCTGATGCTGCCCGGAGGGGCAATGCCGCGTGGCGGCGTATGTGCGAGCACGACATCTCTG	1140
Db	1081	ATGCTGATGTTGTCAGAGAGCGCTCCCGCGCGCGGCTTACCGCGAGCACGACATCTCTG	1140
Qy	1141	ATCGTGGGGAGAGACGGGTGCCGGAACATCAACGGGCTTCGCGTTCCGTCGAGAACACACC	1200
Db	1141	ATCGTGGGGAGAGACGGGCGCAGAGAACATTCACCGATTCCTTCGGGCTTGACACACACC	1200
Qy	1201	ATCATCCGCAAC 1212	
Db	1201	ATCATCCGCAAC 1212	
RESULT 14			
ACCG9521			
XX	ID	ACCG9521 standard; DNA; 1212 BP.	
XX	AC	ACCG9521;	
XX	DT	21-JUL-2003 (first entry)	
XX	DE	Mutant Erwinia creatinase Ctsd7 encoding DNA SEQ ID NO:21.	
XX	XX	Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;	
XX	KW	Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;	
XX	KW	chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;	
XX	XX	gene; ds.	
OS	OS	Erwinia sp.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
FT	CDS	1..1212	
FT	FT	/*tag= a	
FT	FT	/partial	
FT	FT	/EC number= "3.5.3.3"	
FT	FT	/product= "creatinase mutant Ctsd7"	
XX	XX	/note= "no stop codon given"	
XX	XX	EP1298213-A1.	
XX	XX	02-APR-2003.	

XX 17-SEP-2002; 2002EP-00020793.  
 PF Best Local Similarity 87.8%; Pred. No. 7.8e-156;  
 PR 20-SEP-2001; 2001EP-00121780.  
 XX (HOFF) ROCHE DIAGNOSTICS GMBH.  
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.  
 XX

PI Shao Z, Schmuck R, Kratzsch P, Kenklee J, Weisner H;  
 XX WPI; 2003-38334/37.  
 DR P-PSDB; ABR43480.  
 XX

PT New variant of an Erwinia-type creatinase modified relative to a wild-  
 PT type creatinase having creatinase activity, useful for determining  
 PT creatinase and/or creatine concentration in a sample.  
 XX

Example 4; Page 40-42; 51pp; English.

XX The present invention describes a variant of an Erwinia-type creatinase  
 CC (I) modified relative to a wild-type creatinase having creatinase  
 CC activity. The variant comprises at least one amino acid substitution at a  
 CC position of the fully defined 404 amino acid sequence given in ABR43467;  
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has  
 CC the EC number EC 3.5.3.3 and is also known as creatine amidohydrolase.  
 CC Also described is a reagent (II) for determining creatine comprising the  
 CC Erwinia-type creatinase variant. The variant is useful for determining  
 CC creatinine and/or creatine concentration in a sample. Measuring  
 CC creatinine and/or creatine are useful for diagnosing uraemia, chronic  
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some  
 CC other related diseases. The mutant enzymes have improved stability, lower  
 CC conductivity and/or lower Km-values for creatine; they are much better  
 CC suited to detection methods for creatine. The present sequence encodes a  
 CC mutant Erwinia creatinase from the present invention  
 XX

SQ Sequence 1212 BP; 239 A; 391 C; 362 G; 220 T; 0 U; 0 Other;

Query Match 80.5%; Score 975.2; DB 8; Length 1212;

Best Local Similarity 87.8%; Pred. No. 7.8e-156;  
 Matches 1064; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGATGACGACATGTTGACGTTGAAATGACCAACGCGAGAAAGATTATTCGCG 60  
 DB 1 ATGATGACGACATGTTGACGTTGAAATGACCAACGCGAGAAAGATTATTCGCG 60  
 QY 61 TTTTGGATGCCGAGATGACCGCGCCAAACGAGTTCGGCGTGGATGGCCAAAGAC 120  
 DB 61 TTTTCCGATGCCGAGATGACCGCGCCAAACGAGTTCGGCGTGGATGGCCAAAGAC 120  
 QY 121 AATGTCGATGCCGCGCTGTTCACTCTTATCACTGATCACTACTATTCGGCTGG 180  
 DB 121 AATGTCGATGCCGCGCTGTTCACTCTTATCACTGATCACTACTATTCGGCTGG 180  
 QY 121 GACGTGACGATGCGCTGTTCACTCTTATCACTGATCACTACTATTCGGCTGG 180  
 DB 121 GACGTGACGATGCGCTGTTCACTCTTATCACTGATCACTACTATTCGGCTGG 180  
 QY 181 TACTGCTATTTGAGCGAAGTACGAGATGATGATGACCAACGCGCCAGAGATT 240  
 DB 181 TACTGCTATTTGAGCGAAGTACGAGATGATGATGACCAACGCGCCAGAGATT 240  
 QY 241 TCGGCGGATGACG 300  
 DB 241 TCGGCGGATGACG 300  
 QY 301 ACCGACTGGCG 360  
 DB 301 ACCGACTGGCG 360  
 QY 361 CGCATCGGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 420  
 DB 361 CGCATCGGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 420  
 QY 421 CTACCGGCGGCTGACCTTCGATGACATCAGCCGCGATGATGATGATGATGAT 480  
 DB 421 CTGCGCGGCGGCTGACCTTCGATGATCAGCCGCGATGATGATGATGATGAT 480

QY 481 TCGCTGAGAGACGAGACGTGATCCGAGAGCGCCGCGTGTGATGATGATGATGAT 540  
 DB 481 TCGCTGAGAGACGAGACGTGATCCGAGAGCGCCGCGTGTGATGATGATGATGAT 540  
 QY 541 GCGTGGCGGCTGCCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 DB 541 GCGTGGCGGCTGCCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 QY 601 AATGCGATGATCCGCGAGATGCGCCAAATGCTTCCCTTGTGTGAGACTGATGAC 660  
 DB 601 AATGCGATGATCCGCGAGATGCGCCAAATGCTTCCCTTGTGTGAGACTGATGAC 660  
 QY 661 AATGCGATGATCCGCGAGATGCGCCAAATGCTTCCCTTGTGTGAGACTGATGAC 720  
 DB 661 AATGCGATGATCCGCGAGATGCGCCAAATGCTTCCCTTGTGTGAGACTGATGAC 720  
 QY 721 GTGCAATCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 DB 721 GTGCAATCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 QY 781 GCGCTGAGCGCGCGCTTCTGCGGACCATGATGATGATGATGATGATGATGAT 840  
 DB 781 GCGCTGAGCGCGCGCTTCTGCGGACCATGATGATGATGATGATGATGATGAT 840  
 QY 841 AAGAACGTCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 DB 841 AAGAACGTCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 QY 901 GACATCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 901 GACATCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 961 TTTGCGTATGCGCGCTTCTGCGGACCATGATGATGATGATGATGATGATGAT 1020  
 DB 961 TTTGCGTATGCGCGCTTCTGCGGACCATGATGATGATGATGATGATGATGAT 1020  
 QY 1021 GAGCTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 1021 GAGCTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 QY 1081 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 DB 1081 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 QY 1141 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 DB 1141 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 QY 1201 ATCATCGCGAAC 1212  
 DB 1201 ATCATCGCGAAC 1212

RESULT 15  
 ACC69520  
 ID ACC69520 standard; DNA; 1212 BP.  
 XX  
 AC ACC69520;  
 XX  
 DT 21-JUL-2003 (first entry)  
 XX  
 DE Mutant Erwinia creatinase CTsd2 encoding DNA SEQ ID NO:19.  
 XX  
 KW Erwinia; creatinase; creatine amidohydrolase; enzyme; EC 3.5.3.3;  
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;  
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;  
 XX gene; ds.  
 XX  
 OS Erwinia sp.  
 XX  
 OS Synthetic.  
 XX  
 FH Key  
 FT CDS Location/Qualifiers  
 1..1212

```

FT      /*tag= a
FT      /partial
FT      /BC-number= "3.5.3.3"
FT      /product= "creatinase mutant CT6d2"
FT      /note= "no stop codon given"
XX      EP1298213-A1.
XX      02-APR-2003.
XX      17-SEP-2002; 2002EP-00020793.
XX      20-SEP-2001; 2001EP-00121780.
XX      (HOFF) ROCHE DIAGNOSTICS GMBH.
XX      (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX      Shao Z, Schmueck R, Kratzsch P, Kenkiles J, Weisner H;
XX      WPI: 2003-383834/37.
XX      P-PDB; ABR43479.
XX      New variant of an Erwinia-type creatinase modified relative to a wild-
XX      PT type creatinase having creatinase activity, useful for determining
XX      PT creatinase and/or creatine concentration in a sample.
XX      Example 4; Page 37-39; 51pp; English.
XX      The present invention describes a variant of an Erwinia-type creatinase
XX      CC (I) modified relative to a wild-type creatinase having creatinase
XX      CC activity. The variant comprises at least one amino acid substitution at a
XX      CC position of the fully defined 404 amino acid sequence given in ABR43467.
XX      CC these are selected from N130, N203, I278, I1304 and P355. Creatinase has
XX      CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
XX      CC Also described is a reagent (II) for determining creatine concentration
XX      CC Erwinia-type creatinase variant. The variant is useful for determining
XX      CC creatinine and/or creatine concentration in a sample. Measuring
XX      CC creatinine and creatine are useful for diagnosing uremia, chronic
XX      CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
XX      CC other related diseases. The mutant enzymes have improved stability, lower
XX      CC conductivity and/or lower Km-values for creatine; they are much better
XX      CC suited to detection methods for creatine. The present sequence encodes a
XX      CC mutant Erwinia creatinase from the present invention
XX      CC
XX      Sequence 1212 BP; 239 A; 393 C; 360 G; 220 T; 0 U; 0 Other;
XX
Query Match      80.3%; Score 973.6; DB 8; Length 1212;
Best Local Similarity 87.7%; Pred. No. 1.5e-155;
Matches 1063; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

```

```

QY      361 CGCATGCGATGAGTTGACCACTGCAATCTTCGACTTCCGCCCGCAGCTCGAGAGACC 420
DB      361 CGCATGCGATGAGTTGACCACTGCAATCTTCGACTTCCGCCCGCAGCTCGAGAGAGCG 420
QY      421 CTACCGGCGGCTGCACTGCTGCAATGACCGACCTCGATGTGATGCGCAACATCAAG 480
DB      421 CTGCGCGGCGGCTGCACTGCTGCAATGACCGACCTCGATGTGATGCGCAACGCTCAAG 480
QY      481 TCGCTCGAAGACAGAGAGTGTATCCGCGAAGCGCCGCGTGTGACGTCGCGCGCGG 540
DB      481 TCGCTCGAAGACAGAGAGTGTATCCGCGAAGCGCGCGTGTGACGTCGCGCGCGTGC 540
QY      541 GCGTGGCGGCTGCGCATCAAGCGCGCGTGTGCGGACATGAAGTGCATCGCCACAC 600
DB      541 GCGTGGCGGCTGCGCATCAAGCGCGCGTGTGCGGACATGAAGTGCATCGCCACAC 600
QY      601 AATGCGATGATCGCGAGATCGCAATGCTTCCCTTGTGAGAGCTGATGACACCTGG 660
DB      601 AATGCGATGATCGCGAGATCGCAATGCTTCCCTTGTGAGAGCTGATGACACCTGG 660
QY      661 ACCTGGTTCAGTCCGCGCATCAACACCGAGCGGCGCAATCCGATCCAAACCGCATC 720
DB      661 ACCTGGTTCAGTCCGCGCATCAACACCGAGCGGCGCAATCCGATCCAAACCGCATC 720
QY      721 GTGCAATCGGCGGACATCTTTCGCTCAACACTTCCGATGATTTTGGGCTACTACAC 780
DB      721 GTGCAATCGGCGGACATCTTTCGCTCAACACTTCCGATGATTTTGGGCTACTACAC 780
QY      781 GCGCTGAGAGCGGACGCTGCTTGCAGCAGATGATGATGACGCGGCTCGACATCTGGAG 840
DB      781 GCGCTGAGAGCGGACGCTGCTTGCAGCAGATGATGATGACGCGGCTCGACATCTGGAG 840
QY      841 AAGAACGTGGCGGCTGATCGCGCGGCTGAGTGAACCGCGGCGCGCTGGCAAG 900
DB      841 AAGAACGTGGCGGCTGATCGCGCGGCTGAGTGAACCGCGGCGCGCTGGCAAG 900
QY      901 GATATCGGATGAGTGAATCAAGAGTGAACCGGAGTGGAGCTGTAAGTACCGCTCC 960
DB      901 GATATCGGATGAGTGAATCAAGAGTGAACCGGAGTGGAGCTGTAAGTACCGCTCC 960
QY      961 TTGCGCTAGCGGCACTCTTCCGCGGCTGCTTCCACTACTAGGCGCGGCGGCTG 1020
DB      961 TTGCGCTAGCGGCACTCTTCCGCGGCTGCTTCCACTACTAGGCGCGGCGGCTG 1020
QY      1021 GAGCTGCGGAGGACATGACACCGAGTGAAGCCCGGATGCTTCCATGAGGCGG 1080
DB      1021 GAGCTGCGGAGGACATGACACCGAGTGAAGCCCGGATGCTTCCATGAGGCGG 1080
QY      1081 ATGCTGATGCTGCGGAGGAGTGCCTGCGGCGGCTGATCGGAGACAGACATCTCTG 1140
DB      1081 ATGCTGATGCTGCGGAGGAGTGCCTGCGGCGGCTGATCGGAGACAGACATCTCTG 1140
QY      1141 ATCGTGGGAGGAGAGCGTGCAGAGACATCAACCGGCTTCCGTTCCGTCGGAACACAC 1200
DB      1141 ATCGTGGGAGGAGAGCGTGCAGAGACATCAACCGGCTTCCGTTCCGTCGGAACACAC 1200
QY      1201 ATGATCGGCAAC 1212
DB      1201 ATGATCGGCAAC 1212

```

Search completed: November 19, 2004, 16:39:06  
Job time : 664 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 12:40:44 ; Search time 124 Seconds

(without alignments)  
6947.388 Million cell updates/sec

Title: US-10-807-228-2

Perfect score: 1212  
Sequence: 1 ATGACTGACGACATCTTGCA.....AACACACATCATCCGCAAC 1212

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/pdata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/pdata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/pdata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/pdata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/pdata/1/ina/PCUS\_COMB.seq:\*
- 6: /cgn2\_6/pdata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	3	US-08-799-897-2
2	1204	99.3	1215	2	US-08-947-726A-1
3	74	6.1	858	4	US-09-252-991A-4171
4	74	6.1	963	4	US-09-252-991A-4435
5	74	6.1	1176	4	US-09-252-991A-4314
6	64.8	5.3	3157	2	US-08-939-002A-1
7	63.6	5.2	2064	1	US-08-343-428-1
8	62.4	5.1	1294	3	US-09-025-691-2
9	60.6	5.0	1599	4	US-09-475-515-54
10	60.6	5.0	2112	4	US-09-475-515-57
11	60.6	5.0	2112	4	US-09-475-515-56
12	60.6	5.0	2181	4	US-09-475-515-58
13	60.6	5.0	2634	4	US-09-475-515-64
14	59.6	5.0	4766	4	US-09-475-515-73
15	59.6	4.9	2712	3	US-09-025-691-4
16	59.2	4.9	1056	4	US-09-266-965-66
17	59.2	4.9	53500	4	US-09-266-965-76
18	58.4	4.8	1965	4	US-09-220-132-21
19	57.6	4.8	1293	2	US-08-924-440-1
20	56.8	4.7	2943	1	US-08-042-747A-7
21	56.4	4.7	1005	4	US-09-252-991A-1549
22	56.4	4.7	1590	4	US-09-252-991A-1617
23	56.4	4.7	1693	3	US-09-320-878-23
24	56.2	4.6	1693	3	US-09-141-908-19
25	56.2	4.6	1693	4	US-09-657-440-23
26	55.8	4.6	1399	1	US-08-471-033-24
27	55.8	4.6	1399	1	US-08-471-033-24

28	55.8	4.6	1399	2	US-08-471-044-24	Sequence 24, Appl
29	55.8	4.6	1399	2	US-08-463-483A-24	Sequence 24, Appl
30	55.8	4.6	1399	2	US-08-471-046A-24	Sequence 24, Appl
31	55.8	4.6	1399	2	US-08-470-566B-24	Sequence 24, Appl
32	55.8	4.6	1399	2	US-08-469-334-24	Sequence 24, Appl
33	55.8	4.6	1399	3	US-09-300-529-24	Sequence 24, Appl
34	55.6	4.6	1082	4	US-09-881-165-4	Sequence 4, Appl
35	55.4	4.6	2655	1	US-08-471-033-17	Sequence 17, Appl
36	55.4	4.6	2655	1	US-08-471-033-26	Sequence 26, Appl
37	55.4	4.6	2655	2	US-08-471-044-17	Sequence 17, Appl
38	55.4	4.6	2655	2	US-08-471-044-26	Sequence 26, Appl
39	55.4	4.6	2655	2	US-08-463-483A-17	Sequence 17, Appl
40	55.4	4.6	2655	2	US-08-463-483A-26	Sequence 26, Appl
41	55.4	4.6	2655	2	US-08-471-046A-17	Sequence 17, Appl
42	55.4	4.6	2655	2	US-08-471-046A-26	Sequence 26, Appl
43	55.4	4.6	2655	2	US-08-470-566B-17	Sequence 17, Appl
44	55.4	4.6	2655	2	US-08-470-566B-26	Sequence 26, Appl
45	55.4	4.6	2655	2	US-08-469-334-17	Sequence 17, Appl

# ALIGNMENTS

RESULT 1  
US-08-799-897-2  
Sequence 2, Application US/08799897  
Patent No. 6080553  
GENERAL INFORMATION:  
APPLICANT: Sogabe, Ateushi  
APPLICANT: Hattori, Takashi  
APPLICANT: Nishiyama, Yoshiaki  
APPLICANT: Kawamura, Yoshihisa  
TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: Illinois  
COUNTRY: US  
ZIP: 60601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08799,897  
FILING DATE: 13-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 25435/1996  
FILING DATE: 13-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Robert F. Green  
REGISTRATION NUMBER: 27555  
REFERENCE/DOCKET NUMBER: 78064  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
TELEX: 25-3533  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1212 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
ORGANISM: Alcaligenes faecalis  
STRAIN: TE3581 (ERM P-14237)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 1 to 1212  
US-08-799-897-2

Query Match 99.9%; Score 1210.4; DB 3; Length 1212;  
Best Local Similarity 99.9%; Pred. No. 5.1e-255;  
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGACTGAGACATGTTGCACTGATGAATGGCAACAAGGCGAAGAAATTTATTCGCG 60
DB 1 ATGACTGAGACATGTTGCACTGATGAATGGCAACAAGGCGAAGAAATTTATTCGCG 60
QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAACGATTCGGGCGTGAATGGCGAAGAC 120
DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAACGATTCGGGCGTGAATGGCGAAGAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCACTACTATTCCGGCTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTTATCACTGATCACTACTATTCCGGCTGGCTG 180
QY 181 TACTGCTATTTGGAGAGCAAGTACGGCATGTCATGACCAACAACCGCACAGAGATT 240
DB 181 TACTGCTATTTGGAGAGCAAGTACGGCATGTCATGACCAACAACCGCACAGAGATT 240
QY 241 TCGGCGGCGCATGACGCGCGCGCAGCCCTGCGCGCGGAGCTTGGGCGCAACATCACTAC 300
DB 241 TCGGCGGCGCATGACGCGCGCGCAGCCCTGCGCGCGGAGCTTGGGCGCAACATCACTAC 300
QY 301 ACCGACTGGCGCGCGCAATTTTATTCGCGCGCTGCGCACGTCGACGCGCGCGCAAG 360
DB 301 ACCGACTGGCGCGCGCAATTTTATTCGCGCGCTGCGCACGTCGACGCGCGCGCAAG 360
QY 361 CGCATTCGGATGAGTTCGACCACTGCACTTCGCGCGCGCGCACGTCGACGAGCC 420
DB 361 CGCATTCGGATGAGTTCGACCACTGCACTTCGCGCGCGCGCACGTCGACGAGCC 420
QY 421 CTACCGGCGCGCATGCTTCTGCACTCACTGATGATGATGATGATGATGATGATGATG 480
DB 421 CTACCGGCGCGCATGCTTCTGCACTCACTGATGATGATGATGATGATGATGATGATG 480
QY 481 TCGCTCGAAGAGCAGAACTGATTCGCGAAGGCGCGCGCTGATGATGATGATGATGATG 540
DB 481 TCGCTCGAAGAGCAGAACTGATTCGCGAAGGCGCGCGCTGATGATGATGATGATGATG 540
QY 541 GCGTCCGCGCGCTGCGCATCAAGGCGCGCGCTGCGCATGATGATGATGATGATGATG 600
DB 541 GCGTCCGCGCGCTGCGCATCAAGGCGCGCGCTGCGCATGATGATGATGATGATGATG 600
QY 601 AATGCGATGATCGCGGAGATGCGCAATGCTTCCCTTCTGATGATGATGATGATGATG 660
DB 601 AATGCGATGATCGCGGAGATGCGCAATGCTTCCCTTCTGATGATGATGATGATGATG 660
QY 661 ACCTGTTTCAGTTCGCGGATCAACCGGCGCGCGCATTCGCGTCAACCGGCGATC 720
DB 661 ACCTGTTTCAGTTCGCGGATCAACCGGCGCGCGCATTCGCGTCAACCGGCGATC 720
QY 721 GTGCAATCGCGGAGATCTTTCGCTCAACACTTTCGCGATATCTTGGCTACTACACC 780
DB 721 GTGCAATCGCGGAGATCTTTCGCTCAACACTTTCGCGATATCTTGGCTACTACACC 780
QY 781 GCGCTGAGAGCGCAGCTGTTCTGAGACCATGTCGATGACGCGGCTGACATCTGGAG 840
DB 781 GCGCTGAGAGCGCAGCTGTTCTGAGACCATGTCGATGACGCGGCTGACATCTGGAG 840
QY 841 AAGAACTGCGCGCTGATTCGCGCGCGCTGAGCTGATCAACCGCGCGCGCTGCAAG 900
DB 841 AAGAACTGCGCGCTGATTCGCGCGCGCTGAGCTGATCAACCGCGCGCGCTGCAAG 900
QY 901 GACATCGCGATGAGCTCAACAGATGATTCGCGGATGAGGAGCTGCTGAAGTACCGCTCC 960
DB 901 GACATCGCGATGAGCTCAACAGATGATTCGCGGATGAGGAGCTGCTGAAGTACCGCTCC 960
QY 961 TTGGGCTATGAGCACTCTTGGGCGCTGCTGCACTACTAGTTCGAGGCTGGCGCGTG 1020
```

```
DB 961 TTGGGCTATGAGCACTCTTGGGCGCTGCTGCACTACTAGTTCGAGGCTGGCGCGTG 1020
QY 1021 GAGTTCGCGAGACATGACACCGAGCTGAAGCCCGCATGCTGATCTCAATGAGCGG 1080
DB 1021 GAGTTCGCGAGACATGACACCGAGCTGAAGCCCGCATGCTGATCTCAATGAGCGG 1080
QY 1081 ATGTCGATGCTGCGGAGGCGCATGCGCGGCTATGCGGAGCAACATCTCTG 1140
DB 1081 ATGTCGATGCTGCGGAGGCGCATGCGCGGCTATGCGGAGCAACATCTCTG 1140
QY 1141 ATGTCGCGGAGAGACGCTGCGGAGCAATCAACCGGCTTCCGTTGCTCGGAAACAC 1200
DB 1141 ATGTCGCGGAGAGACGCTGCGGAGCAATCAACCGGCTTCCGTTGCTCGGAAACAC 1200
QY 1201 ATCATCGGCAC 1212
DB 1201 ATCATCGGCAC 1212
```

```
RESULT 2
US-08-947-726A-1
; Sequence 1, Application US/08947726A
; Patent No. 5932466
; GENERAL INFORMATION:
; APPLICANT: Furukawa, Keisuke
; APPLICANT: Ichikawa, Toshio
; APPLICANT: Suzuki, Masaru
; APPLICANT: Koyama, Yasuji
; TITLE OF INVENTION: A NOVEL CREATINE AMIDINOHYDROLASE GENE,
; TITLE OF INVENTION: A NOVEL RECOMBINANT DNA, AND A PROCESS FOR PRODUCING
; NUMBER OF INVENTIONS: 5
; CORRESPONDENCE ADDRESS:
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/947,726A
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/535,444
; FILING DATE: 27-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 94/235737
; FILING DATE: 29-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence III, Stanton T.
; REGISTRATION NUMBER: 25,736
; REFERENCE/DOCKET NUMBER: 7005-120-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes sp. KS-85
; STRAIN: FERM BP-4487
```



US-08-947-726A-1

Query Match 99.3%; Score 1204; DB 2; Length 1215;  
Best Local Similarity 99.6%; Pred. No. 1.3e-253;  
Matches 1207; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATGACGACATGTTGCACTGATGAAATGACACAAAGGCGGAGAAATTTTGGCCG 60  
DB 1 ATACATGACACATGTTGCACTGATGAAATGACACAAAGGCGGAGAAATTTTGGCCG 60  
QY TTTTCGAGATGCGAGATGACCCCGCGCAAAAGAGCTTGGCGGTGATGCGCAAGAC 120  
DB TTTTCGAGATGCGAGATGACCCCGCGCAAAAGAGCTTGGCGGTGATGCGCAAGAC 120  
QY 121 AATGTGATGCGCGCTGTTCACTTATCACTGATCACTAATTTCCGCTGGCTG 180  
DB 121 AATGTGATGCGCGCTGTTCACTTATCACTGATCACTAATTTCCGCTGGCTG 180  
QY 181 TACTGCTATTTGGAAGCAAGTACGGCATGTCATGACCAACAAACGCGACAGAGATT 240  
DB 181 TACTGCTATTTGGAAGCAAGTACGGCATGTCATGACCAACAAACGCGACAGAGATT 240  
QY 241 TCGCGCGGATGACGCGCGCGCACCCCTGCGCGCGACGTTCCGCGCAACATCACTAC 300  
DB 241 TCGCGCGGATGACGCGCGCGCACCCCTGCGCGCGACGTTCCGCGCAACATCACTAC 300  
QY 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGTGGCGCACTGACACAGGCGCGAAG 360  
DB 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGTGGCGCACTGACACAGGCGCGAAG 360  
QY 361 CGCATGCGGATGAGTTGCAACGTCATCTGACTCCGCGCGCAAGCTCGAGAGAGCC 420  
DB 361 CGCATGCGGATGAGTTGCAACGTCATCTGACTCCGCGCGCAAGCTCGAGAGAGCC 420  
QY 421 CTACCGCGGCTGCACTTCTGTCATCACTGACCAAGCTTCTGATGATGCGCAACATCAAG 480  
DB 421 CTACCGCGGCTGCACTTCTGTCATCACTGACCAAGCTTCTGATGATGCGCAACATCAAG 480  
QY 481 TCGCTCGAAGAGCAAGCTGATCCCGAAGGCGCGCGTGTGATCGTGGCGCGCG 540  
DB 481 TCGCTCGAAGAGCAAGCTGATCCCGAAGGCGCGCGTGTGATCGTGGCGCGCG 540  
QY 541 GCGTGGCGGCTGCACTCAAGGCGCGCGCGCGCAACAAATGCGCATCGCGACACAC 600  
DB 541 GCGTGGCGGCTGCACTCAAGGCGCGCGCGCGCAACAAATGCGCATCGCGACACAC 600  
QY 601 AATGCGATGATCGCGAGATCGCAATCGTTCCCTTCTGATGATGATGATGATGATG 660  
DB 601 AATGCGATGATCGCGAGATCGCAATCGTTCCCTTCTGATGATGATGATGATGATG 660  
QY 661 ACTGCTTCAGTGGGCAATCAACGAGGCGCGCAACATCGGTCACCAACCGCATC 720  
DB 661 ACTGCTTCAGTGGGCAATCAACGAGGCGCGCAACATCGGTCACCAACCGCATC 720  
QY 721 GTGCAATCGCGGCACTCTTTGCTCAACCTTCCCATGATCTTGGCTACTACAC 780  
DB 721 GTGCAATCGCGGCACTCTTTGCTCAACCTTCCCATGATCTTGGCTACTACAC 780  
QY 781 GCGCTGAGAGCGAGCTGTTCTGCAACATGTCGATGACGCGCTGCACTCTGGAG 840  
DB 781 GCGCTGAGAGCGAGCTGTTCTGCAACATGTCGATGACGCGCTGCACTCTGGAG 840  
QY 841 AAGAACGTGCGCTGATCGCGCGGCTGCACTGATCAAGCGCGCGCGCTGCAAG 900  
DB 841 AAGAACGTGCGCTGATCGCGCGGCTGCACTGATCAAGCGCGCGCGCTGCAAG 900  
QY 901 GACATGCGCATGAGCTCAACGAGATGACCGGATGAGGAGCTGCTGAAGTACCGCTCC 960  
DB 901 GACATGCGCATGAGCTCAACGAGATGACCGGATGAGGAGCTGCTGAAGTACCGCTCC 960  
QY 961 TTGCGCTATGAGCACTCTTGGCGCTGCTGTCATCACTGATGCGAGGCGCGGCTG 1020  
DB 961 TTGCGCTATGAGCACTCTTGGCGCTGCTGTCATCACTGATGCGAGGCGCGGCTG 1020

QY 1021 GAGCTCGCGAGAGACATCGACACCGAGCTGAAGCCCGGCAATGATGCTTCCATGAGACCG 1080  
DB 1021 GAGCTCGCGAGAGACATCGACACCGAGCTGAAGCCCGGCAATGATGCTTCCATGAGACCG 1080  
QY 1081 AATGTGATGCTGCGGAGGAGGAGATGCGCGGTCGCGGCGGCTATGCGAGCAACATCTCTG 1140  
DB 1081 AATGTGATGCTGCGGAGGAGGAGATGCGCGGTCGCGGCGGCTATGCGAGCAACATCTCTG 1140  
QY 1141 ATGCTGCGGAGAGAGCGGTCCGAGAACATCACCGCTTCCGTTCCGTCGGAACACACC 1200  
DB 1141 ATGCTGCGGAGAGAGCGGTCCGAGAACATCACCGCTTCCGTTCCGTCGGAACACACC 1200  
QY 1201 ATCATTCGCGAAC 1212  
DB 1201 ATCATTCGCGAAC 1212

RESULT 3  
US-09-252-991A-4171  
; Sequence 4171, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4171  
; LENGTH: 858  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4171

Query Match 6.1%; Score 74; DB 4; Length 858;  
Best Local Similarity 46.3%; Pred. No. 2.5e-07;  
Matches 316; Conservative 0; Mismatches 360; Indels 6; Gaps 2;

QY 471 CACCATCAAGTGTCTGGAAGAGCAAGCTGATCCGCAAGGCGCGCTGTGTACCT 530  
DB 81 CACCATCAAGGCGCGAGATGATGAGAAATGCGCATCGCGCGCTGCGCGCGCA 140  
QY 531 CGGCGCGGCGCTGCGCGCTGCGCATCAAGGCGCGCGCTGCGCGCGCAATGCGCAT 590  
DB 141 AGTGTGGAATGATCGGGAACATCAAGCTCGGCGTGAACCGAGAGCTGATCG 200  
QY 591 CGCGCAACCAATGAGATGATCGGAGATCGCAATGCTTCCCTTCTGAGAGCTGAT 650  
DB 201 CATCTGCCACGATATCTGATCAAGAGAGAGAGAGATTCCTGCGCGCGCGCAAC 260  
QY 651 GGAACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 710  
DB 261 GGGCTTTCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 320  
QY 711 CAACCGCATGCTGCAATCGCGGAGATCTTTCGCTCAACACTTCCGATATCTTCCG 770  
DB 321 CGAAGAGCGCTGAGAGAGAGAGAGATCTTCAAGCTGCAATCAAGCTGATCAAGAGAG 380  
QY 771 CTACTAACCGCGCTGAG 830  
DB 381 CTACTAACCGCGAG 440  
QY 831 CATCTGAG 890  
DB 441 CTTCTGCAAGATCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 500  
QY 891 GCGCTGCAAGAGATCGCATGAGCTCAAGAGATGATCGCGAGAGAGAGAGAGAGAGAG 950

Db 501 GCACCTGGGAGATGCG---GCCAATCATCCAGAACCGGAAAGAAAGCCTTCTC 557  
Qy 951 GTACCCCTCTTCCGCTAGTGGCCATCTTCCTGGCGTGTGCTGCACTACTACGCTGCGA 1010  
Db 558 GGTGTCTCCGGAATAGCGGGCCATGCGCATCGGCAAGGTGTCC---ACGAGAGCCGCA 614  
Qy 1011 GCGCGGCGTGGACTCGCCAGACATCGACACCGAGCTGAACCCCGGCACTGTGTCTC 1070  
Db 615 GGTCTCTCACTAGGTGTGTGCGGATACCGGCATCGAGTGAAGAGGACATGATCTTCAC 674  
Qy 1071 CATGAGCCGATGTGATGTGCTGCGGAGGCGATGCCGCTGCGGCGCTATCGCAGCA 1130  
Db 675 CATCGAGCCGATATCAACAGGCGCGCCCGAGAGACCCGCTGTGCGGACGCTGAGC 734  
Qy 1131 CGACATCTGTATGTGTGGGAG 1152  
Db 735 CGCGATCACCAGAGACCGCAG 756

RESULT 4  
US-09-252-991A-4435/c  
; Sequence 4435, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4435  
; LENGTH: 963  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4435

Query Match 6.1%; Score 74; DB 4; Length 963;  
Best Local Similarity 46.3%; Pred. No. 2.5e-07;  
Matches 316; Conservative 0; Mismatches 360; Indels 6; Gaps 2;  
Qy 471 CACCATCAAGTGTGCTGGAAGACGAGATGATCCGGAAGCGCCGCGTGTGACGT 530  
Db 922 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGGCGCGCA 863  
Qy 531 CCGCGCGCGCGCTGCGCGCTGCGCATCAAGCGCGCGTCCCGAGCATGAAGTGGCAT 590  
Db 862 AGTGTGGAATGATGTGGGAGACATCAAGCGCGCGTGAACCGAGAGCTGATCG 803  
Qy 591 CGGCACACCAATGCGATATCGCGAGATCGCCAATGTTCCCTGTGTGAGTGTAT 650  
Db 802 CATGTCCACGATATATGTCAACGAGAGGAGATTCCTCGCCCTGAACTACAA 743  
Qy 651 GGAACCTGAGCTGTGCTCAAGTGTGGGATCAACCGAGCGCGCAGCAATCCGTAC 710  
Db 742 GGGCTTCCCAAGTGTATGCACTGTGATCAACATGTGTGTGCTGATGGCATGCCAA 683  
Qy 711 CAACCGCATGTGCAATCCGCGGACATCTTTCCTCAACACTTCCCATGATCTTGG 770  
Db 682 CGAGAGCGCTGAAGAGAGGCGACATCTCAACGTGACATCAACGTGATCAAGAGCG 623  
Qy 771 CTACTACACCGCGCTGAGGCGACGCTGTTCGCAACATGTGATGAGCGACCTGCA 830  
Db 622 CTACCAACCGCGGACGACGAAAGATGTTCTGTGTGCAAGATCCCGAATGGCGGACG 563  
Qy 831 CATTTGGAGAGAAAGTGTGCGCTGTGATCGCGCGGCTGAGCTGTATCAAGCGGCGC 890  
Db 562 CCTGTGCGAGATCAACCAAGAAATGATGTAACAAGGCAATTTCCGTGTGCGTCCGCGCGC 503

Qy 891 GCGCTGCAAGACATGTCCATGAGCTCAACGAGATGTACCGGAGTGGACCTGCTGA 950  
Db 502 GCACCTGGGCGATATCG---GCCAATCATCCAGAACCGGAAAGAAAGCCTTCTC 446  
Qy 951 GTACCCCTCTTCCGCTATGAGCCATCTTCGCGGCTGTGCTGCACTACTACGCTGCGA 1010  
Db 445 GGTGTCTCCGGAATAGCGGGCCATGCGCATCGGCAAGGTGTCC---ACGAGAGCCGCA 389  
Qy 1011 GCGCGGCGTGGACTCGCCGAGACATCGACACCGAGCTGAAGCCCGGCACTGTGTCTC 1070  
Db 388 GGTCTCTCACTAGGTGTGTGCGGATACCGGCATCGAGCTGAAGAGGACATGATCTTCAC 329  
Qy 1071 CATGAGCCGATGTGATGTGCTGCGGAGGCGATGCCGCTGCGGCGCTATGCGCAGCA 1130  
Db 328 CATCGAGCCGATGATCAACAGGCGCGCCCGAGAGACCCGCTGTGCGGACGCTGAGC 269  
Qy 1131 CGACATCTGTATGTGTGGGAG 1152  
Db 268 CGCGATCACCAGAGACCGCAG 247

RESULT 5  
US-09-252-991A-4314  
; Sequence 4314, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4314  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4314

Query Match 6.1%; Score 74; DB 4; Length 1176;  
Best Local Similarity 46.3%; Pred. No. 2.6e-07;  
Matches 316; Conservative 0; Mismatches 360; Indels 6; Gaps 2;  
Qy 471 CACCATCAAGTGTGCTGGAAGACGAGATGATCCGGAAGCGCCGCGTGTGACGT 530  
Db 244 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGGCGCGCA 303  
Qy 531 CCGCGCGCGCGCTGCGCGCTGCGCATCAAGCGCGCGTCCCGAGCATGAAGTGGCAT 590  
Db 304 AGTGTGGAATGATGTGGGAGACATCAAGCGCGCGGTGACCAACGAGAGCTGATCG 363  
Qy 591 CGGCACACCAATGCGATATCGCGAGATCGCCAAATGTTCCCTGTGTGAGTGTAT 650  
Db 364 CATGTGCAACGATATATGTCAACGAGAGCGAGATTCCTCGCCCTGAACTACAA 423  
Qy 651 GGAACCTGAGCTGTGCTCAAGTGTGGGATCAACCGAGCGCGGCAACATCCGTAC 710  
Db 424 GGGCTTCCCAAGTGTATGCACTGTGATCAACATGTGTGTGCTGATGGCATGCCAA 483  
Qy 711 CAACCGCATGTGCAATCCGCGGACATCTTTCCTCAACACTTCCCATGATCTTGG 770  
Db 484 CGAGAGCGCTGAAGAGAGGAGCATCTCAAGTGTGACATCAACGTGATCAAGAGCG 543  
Qy 771 CTACTACACCGCGCTGAGCGGACGCTGTTCGCAACATGTGATGAGCGACGCTGCA 830  
Db 544 CTACCAACCGCGGACGACCAAGATGTTCTGTGTGCGCAAGACCCCGGATGGCGGACG 603  
Qy 831 CATTTGGAGAGAAAGTGTGCGCTGTGATCGCGCGGCTGAGCTGTATCAACCGGCGC 890

Db 604 CCTCTGCAGATACCCAGAGATGATGTAAGAGGATTTCCGTGTGCTCCGGCCG 663  
Qy 891 GCGCTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950  
Db 664 GCACTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 951 GTACCGCTCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1010  
Db 721 GGTGTCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
Qy 1011 GCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1070  
Db 778 GGTCTCTCACTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837  
Qy 1071 CATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130  
Db 838 CATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 897  
Qy 1131 CGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152  
Db 898 CGGATCACCAGAGACCGCAG 919

RESULT 6  
US-08-939-002A-1

Sequence 1, Application US/08939002A  
Patent No. 5849529

GENERAL INFORMATION:  
APPLICANT: HAYASHI, KIYOSHI  
APPLICANT: LIU, AIMIN  
APPLICANT: LI, HEBIAO  
APPLICANT: HARAGUCHI, KAZUTOMO  
APPLICANT: KITAHARA, YOSHIKI  
TITLE OF INVENTION: CELLOBIOSE PHOSPHORYLASE GENE, VECTOR  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/939, 002A  
FILING DATE: 26-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 221193/1997  
FILING DATE: 04-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION/DOCKET NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 8361-001-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3157 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Cellvibrrio gilvus

STRAIN: ATCC 13127  
INDIVIDUAL ISOLATE: Direct Origin: puc-2  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 359..2824  
OTHER INFORMATION: /note="METHOD FOR DETERMINING  
OTHER INFORMATION: SEQUENCE: E"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 359..2824  
US-08-939-002A-1  
Query Match 5.3%; Score 64.8; DB 2; Length 3157;  
Best Local Similarity 42.1%; Pred. No. 3.3e-05;  
Matches 436; Conservative 0; Mismatches 597; Indels 3; Gaps 1;  
Qy 180 GTACTGTATTTTCGACGCAATGATGATGATGATGATGATGATGATGATGAT 239  
Db 802 GTTCGTGAGTTCTGCTGTGAGACGGCAGAGACGACGACGACGACGACGACGAC 861  
Qy 240 TTCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299  
Db 862 GTGATCGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 921  
Qy 300 CACGACATGCGCGCGGACCAATTTTATGCGCGCGGCGCGGCGCGGCGCGGCGCA 359  
Db 922 GACCGAGTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 981  
Qy 360 GCGCATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419  
Db 982 CGGCTTCGACGACGACGCGGACGACGACGACGACGACGACGACGACGACGACGAC 1041  
Qy 420 CTTACCGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
Db 1042 CTTCCCGCGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101  
Qy 480 GTGCTGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
Db 1102 GCACTTCCTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161  
Qy 540 GCGCTGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 599  
Db 1162 CTACCTGAG 1221  
Qy 600 CAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659  
Db 1222 GCGCGCGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281  
Qy 660 GACCTGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
Db 1282 GAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341  
Qy 720 CTTGCAATCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
Db 1342 CGAGAGCTGAG 1401  
Qy 780 GCGCTGAG 839  
Db 1402 CATGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1461  
Qy 840 GAG 896  
Db 1462 CTCCACAGAT 1521  
Qy 897 CAAGAGATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956  
Db 1522 CATGACATGCGCTGAT 1581  
Qy 957 CTCTTGGCTATGAG 1016  
Db 1582 CACGAT 1641  
Qy 1017 CGTGAGCTGCGGAG 1076

Db 1642 CGCGGGGCTGGCGCGATCAAGAGTCCGGGACTGGGGATCTCTGAGAGAGCCCG 1701  
Qy 1077 GCGGATGATGATCTGCGGAGGCGATGCGCGGCGGCTATGCGAGACGACAT 1136  
Db 1702 GCGGTCGACAAAGAGCCCGCTCCGAGGTCCGCTGTGACACCTGACGCGCTCTT 1761  
Qy 1137 CCGTATGTCGGGAGGAGCGTCCGAGACATCACCGGCTCCCGTGGTCCGGAACA 1196  
Db 1762 CCAAGTCAGGTCGACAAACCGCGCCGCGACGCGCTGCGCTCATGAGCGCTCCACTG 1821  
Qy 1197 CACCATCATTCGCAAC 1212  
Db 1822 GAACGACTGCTCAAC 1837

RESULT 7  
US-08-343-428-1  
; Sequence 1, Application US/08343428  
; Patent No. 5665586  
; GENERAL INFORMATION:  
; APPLICANT: Nakamura, Etano  
; APPLICANT: Tezuka, Hirohige  
; APPLICANT: Kitadokoro, Kengo  
; APPLICANT: Shin, Naeanu  
; APPLICANT: Teraoka, Hiroshi  
; TITLE OF INVENTION: No. 5665586el Protease  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720KB storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS Dos 5.0  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/343,428  
; FILING DATE: 18-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/793/00592  
; FILING DATE: 30-APR-1993  
; APPLICATION NUMBER: JAPAN 4-126511  
; FILING DATE: 19-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley, Jr., James F.  
; REGISTRATION NUMBER: 27794  
; REFERENCE/DOCKET NUMBER: SHGN-7  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)596-9000  
; TELEFAX: (212)596-9090  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2064  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Streptomyces fradiae  
; FEATURE:  
; NAME/KEY: -35 signal  
; LOCATION: 359..364  
; IDENTIFICATION METHOD: by experiment  
; NAME/KEY: -10 signal  
; LOCATION: 378..383  
; IDENTIFICATION METHOD: by experiment  
; NAME/KEY: CDS

; LOCATION: 435..1505  
; IDENTIFICATION METHOD: by experiment  
; NAME/KEY: sig peptide  
; LOCATION: 435..944  
; IDENTIFICATION METHOD: by experiment  
US-08-343-428-1

Query Match 5.2%; Score 63.6; DB 1; Length 2064;  
Best Local Similarity 44.5%; Pred. No. 5.5e-05;  
Matches 343; Conservative 0; Mismatches 419; Indels 8; Gaps 2;

Qy 103 GGCTGATGCGCCAAAGACATGTCATGGGCGCTGTTACCTTTATGATGATCAAC 162  
Db 993 GCCTTCAAGTCACCAAGAGCGGTCGTGCTTCTGACCGCGGCGACGACCAAC 1052  
Qy 163 TACTATTCGGCTGCTGCTACTGCTATTTCGAGCCAGTAGATGTCATGACAC 222  
Db 1053 CTCTGCTCAGCTGTGCTGCTCAC--TCCGCGGACAGTCATGCGCTCCGAGGCG 1109  
Qy 223 AACAAAGCAGCAGATTTCCGCGGATTCGAGCGGCGGCGAGCCCTGCGCGCAGCTTC 282  
Db 1110 ACCAGCTTCCGACCAAGACGATGAGCATGCTGCTACAGACCAACGACGTCGAC 1169  
Qy 283 GCGCAACATCACCCTAGACCGACTGGGCGCGGCAATTTCTATCGCGCGCGCCAG 342  
Db 1170 GCGCGGATGACCTGTACAGGCGGCTACAGAGACATGCGCTCGCGGCGGACGCGCTC 1229  
Qy 343 CTGACCAAGGCGCGCAAGCGCATGCGCATGAGTTGACACAGTCATTCGATTCGCC 402  
Db 1230 GTGGGCGCAGGCGCATGAAGGCGCTCCAGACCAAGTCACAGCGGCGCACGCTCAGC 1289  
Qy 403 CGCCAGCTCGAGAGAGCCCTACCGGCGGTCGATTCGTGACATGACGCGCTGTATG 462  
Db 1290 GCGGTCAAGTCACCTGACCTACACGACAGCGCCCGTTCAGCATGCTCCGACGACC 1349  
Qy 463 TGGATGCGCACCATCAAGTCGCTCGAAGAGCAAGTATCCGGAAGCGCGCGGTG 522  
Db 1350 GCGTCTCGCGCGCGGCGGAGACAGGCGGCGCGCCACTTGGCGGTTCGTCGCGCGC 1409  
Qy 523 TGTGAGTCGCGCGCGCGCGCTGCGCGCGCTGCGCATCAAGCGCGGTGCCGACATGA 582  
Db 1410 ATCCACTCGGAGAGCTCCGCGTGCACGCGGCAACAGCGCTCCGCTCCATCCACGCGGTC 1469  
Qy 583 GTGGGATGCGCACCAACCATGTCATGTCGCGGATGCGCAATGTCCTTCGTCG 642  
Db 1470 GCGGAGCGCTGAGCGCTTATGCGCTGAAC-----GTGATCTGACCGCGCCCGCGC 1524  
Qy 643 GAGCTGATGACACCTGAGCTGTTCCAGTCGCGCATCAACACGAGCGCGCGCAAT 702  
Db 1525 GGGCGGAGTGGCAGAGTGGGGGTATGATGATGTCTCCATACCGCCACACTACAG 1584  
Qy 703 CCGGTACCAACCGCATGTCGATCCGCGCATCTCTTTGCTCAACCTTCCGATG 762  
Db 1585 CTGGCACCAGCGCTGCTGCTGCTGATCCGCGCTCTCTCCAGAGAGCGCTTCGAGCGGAC 1644  
Qy 763 ATCTTCGGCTACTACAGCGCGTGGAGGACGCTGTTCTGAGCATGTCGATGAGCGC 822  
Db 1645 CTGAGCGAGAGAGATGGAGACGCGCTCGCGCGCTTCACGCGCTTACCGGAGACCGC 1704  
Qy 823 AGCTTCGACATCTGGAGAGAAAGCGTGCATGTCGCGCGCGGCTGCA 872  
Db 1705 GCGCGGCGGATGTGGCGCAGGCGAGCGTCTGATGAGCGCGCGGTGTCA 1754

RESULT 8  
US-09-025-691-2  
; Sequence 2, Application US/09025691  
; Patent No. 6069289  
; GENERAL INFORMATION:  
; APPLICANT: Broadway, Roxanne M.  
; APPLICANT: Hartman, Gary E.  
; TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH  
; TITLE OF INVENTION: CHITINOLYTIC ENZYMES



QY 373 GAGTTGACCACTCATCTCAATCCGCGCCGACGTCGAGAAAGCCCTACCGGGCGTC 432  
DB 544 GTGCGGACAAAGGTGAGAGAGTACAGCTTTCTTCAAGCTGAGCGTGGCCATC 603  
QY 433 GACTTGTGACATCAGCCAGCCCTGATGTGATGGGACCATCAAGTGGCTGAAAG 492  
DB 604 GACACGACAAAGCCGCTACCGCTGATCACTGCAACCAAGCTGATCAACCGGCC 663  
QY 493 CAGAACTGATCCGCAAGCGCCGCTGATGAGCTGAGCGCGCGCTGCGGAGCT 552  
DB 664 TGCCCAAGGTGAGTTGAGCCCATCCCATCACTGAGCGCCCGCGCTTCCGC 723  
QY 553 GCCATCAAGCGCGCGCTGCGCGAGTGAAGTGGCATGCCACCAATGAGTATC 612  
DB 724 ATCTCAAGTGAAGACAAAGATTCAACCGGACCGCGCCCTGCAAGAACGTGACACC 783  
QY 613 CGGAGATGCGCAATCGTTCCCTTGTGAGAGCTGATGGAACCTGGAACCTGTTCCAG 672  
DB 784 GTGCAATGACCAACCGGATCCGCGCTGATGAGACCAAGCTGCTGGAACGGCAGC 843  
QY 673 TGGGAGATCAACACCGCGCGGCAATCCGATCAACCAAGCATGTCATCCGCGC 732  
DB 844 CTGGCGAGAGAGATCGTGTGCGCTCCGAGAACCTTCAACGCAACGCGCAAGACATC 903  
QY 733 GACATCTTTGCTCAACACTTCCGATGATCTTGGCTACTACCGCGCTGAGCGC 792  
DB 904 ATCTGACAGCTGAACAGTCCGTGAGATCACTGATCCGCGCCCAACCAACACGCGCT 963  
QY 793 AGCTGTTCTGACCATGTCATGATGAGCGGAGCTGACATCTGGAGAGAAAGTGGCC 852  
DB 964 AAGAGCATCACTCGAGCCCGCGCGCTTCTACGCGACCGGCGACATCATCGCGAC 1023  
QY 853 GTGCATCGCGCGCGCTCGAGCTGATCAAGCGCGCGCGCTGCAAGACATCGCATC 912  
DB 1024 ATCCGCAAGGCCACTGCAACATCAAGCAAGGCCAATGAGCAACCCCTGAGCAGATC 1083  
QY 913 GAGCTCAAGAGATGATCCGCGAGTGGAGACTGCTGAATCCGCTCTTGGCTATGCG 972  
DB 1084 G---TGAGAAAGTGGCGAGGATTCGCGCAACAAAGCAATCTTCTTCAACAGCAGC 1140  
QY 973 CACTCTTGGCGGTGCTGTGCACTACAGCTGCGGAGCGCGCTGAGCGCGAG 1032  
DB 1141 AGCGGCGGAGACCCGAGATCGTGTTCACAGCTTCACTGCGGCGCGAGTTCTTCTAC 1200  
QY 1033 GACATGACACCGAGCTGAAGCCCGGATGTGTCTTCATGAGCGCGATGTGATCTG 1092  
DB 1201 TCGAACACAGCGAGCTTTCAACAGCACTGGAATCATCAAGAGAGTGAACAGAC 1260  
QY 1093 CCGAGAGGATGCGCGGTGCGCGGCTATCGGAGCAGCATCTCTGATCGTGGGAG 1152  
DB 1261 AAGGAGAAAGACACATCTCTGCGCTGCGCATCGCGAGATCTATCAACATGTGGCAG 1320  
QY 1153 GAGCGTGGCGAGAACTACCGGCTTCCGTTGAGTCCGGAACACACATTCATCCGCA 1211  
DB 1321 GAGGTGGGAGAGGCAATGACGCCCCCCCCCATCCGCGCGCAATCAAGTGCAGACAA 1379

RESULT 10  
US-09-475-515-56  
Sequence 56, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 56  
LENGTH: 2112  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: gp140.modus4  
US-09-475-515-56

Query Match 5.0%; Score 60.6; DB 4; Length 2112;  
Best Local Similarity 42.6%; Pred. No. 0.00025;  
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CGCAAGTACGGGATGATGATGACCAACCAAGCGCAGAGATTTCGGCGGCAATCGAC 255  
DB 364 CCCCTGTGTGACCTTGAATCTGACCGAAGAGCTGACCGGACCAACCAAGCGCAGC 423  
QY 256 CGCGGCGAGCCCTGCGCGCGAGCTTGGAGCAACATCACTTACCGGCTGAGCGCGC 315  
DB 424 AGCACAAGGAGCAACAGACAGACGCGGACCAACAGACCAAGCAACAGACCGAC 483  
QY 316 GACAAATTTCTATCGCGCGGTGCGCACT--GACCAAGGCGCGCAAGCGCATTCG 372  
DB 484 AGCTGGAGAAAGATCCCGAGGCGAGATCAAGAACTGACGTTCAACATCAACACAGC 543  
QY 373 GAGTTGACCACTCATCTCAATCCGCGCCGACGTCGAGAAAGCCCTACCGGGCGTC 432  
DB 544 GTGCGGACAAAGGTGAGAGAGTACAGCTTTCTTCAAGCTGAGCGTGGCCATC 603  
QY 433 GACTTGTGACATCAGCCAGCCCTGATGTGATGGGACCATCAAGTGGCTGAAAG 492  
DB 604 GACACGACAAAGCCGCTACCGCTGATCACTGCAACACAGGATGATCAACCGGCC 663  
QY 493 CAGAACTGATCCGCAAGCGCCGCTGATGAGCTGAGCGCGCGCTGCGGAGCT 552  
DB 664 TGCCCAAGGTGAGTGTGAGCGCATCCCATCTCACTGCGCGCCCGCGCTTCCGC 723  
QY 553 GCCATCAAGCGCGGTGCGGAGATGAAGTGGAGTGGCAACCAAGCATGATGATC 612  
DB 724 ATCTGAAGTGAAGAGCAAGAAATTCAACGCGACCGCGCTTGAAGACTGAGCAGC 783  
QY 613 CGCGAGATCGCAATTCGTTCCCTTGTGAGCTGATGAGCACTGCACTGATTCAG 672  
DB 784 GTGCAATGACCAACCGGATCCGCGCTGATGAGACCAAGCTGCTGTAACGGCAGC 843  
QY 673 TGGGATCAACACCGGCGCGCAATCCGATCAACCAAGCATGTCATTCGCGC 732  
DB 844 CTGGCGAGAGAGATGTCGTGCTGCGTCCGAACTTCAACGCAACAGCGCAAGCATC 903  
QY 733 GACATCTTTGCTGCAACACTTCCGATGATCTTGGGATATACACCGCGCTGAGCGC 792  
DB 904 ATCTGACAGCTGAAGAGTCTGTGAGATCACTCATCTGCTCCCAACCAACAGCGCT 963  
QY 793 AGCTGTTCTGACCATGTCATGAGCGCGAGCTTGAATCTGAGAGAAAGTGGCC 852  
DB 964 AAGAGCATCACTCGAGCCCGCGCGCTTCTACGCGACCGGCGAGCATCATTCGGCAGC 1023  
QY 853 GTGCATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAGACATTCGCGATC 912  
DB 1024 ATCCCGCAGCGCACTGCAACATCAAGAGCGCAATGAGCAACCTTCAGAGAGATC 1083  
QY 913 GAGCTCAAGAGATGATCCGAGTGGAGCTGAGTACCGCTCTTGGCTATGAGC 972  
DB 1084 G---TGAGAAAGTGGCGAGGATTCGCGCAACAAAGACCATATATCTTCAACAGCAGC 1140  
QY 973 CACTCTTGGCGGTGCTGCACTACTACGCTGCGAGCGCGCGGTGAGAGCTGCGGAG 1032  
DB 1141 AGCGGCGGAGACCCGAGATGATGTTCACAGCTTCACTGCGGCGGAGATTTCTTCTAC 1200



QY 1033 GACATGACACCGAGTGAAGCCCGCATGTGTCTTCATGAGCCGATGTGATGCTG 1092  
DB 1201 TGCACACACGACGAGTGTTCACACGACCTGGACATCAAGAGAGTGAACAAGACC 1260  
QY 1093 CCGAGAGGATGCCCCGATGCGGCGGCTATCGGAGACAGACATCTGTATCGGAGAG 1152  
DB 1261 AAGAGAAAGACACCATCTCTGCGCTCGCATCGGACATCTCAACATGTGGCAG 1320  
QY 1153 GACGCTGCGGAGACATCAACGGGCTTCCGTTGCTCGGAAACACACCATTCGCA 1211  
DB 1321 GAGTGGGCAAGGCGCATGTACGCCCCCATTCGCGGCGACATCAAGTGACAGCA 1379

## RESULT 11

US-09-475-515-57  
Sequence 57, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 57  
LENGTH: 2112  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-09-475-515-57

Query Match 5.0%; Score 60.6; DB 4; Length 2112;  
Best Local Similarity 42.6%; Pred. No. 0.00025;  
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CCGAAGTACGGATGTCTACCAACCAAGCCACAGCATTTCCGCGGATCCAG 255  
DB 364 CCCCTGTGTGATACCTGTGAATGCAACGCAAGCTGACCGGACCAACGCAAC 423  
QY 256 GCGGCGACAGCCTGCGCGCGGAGCTTCCGAGCAACATCACTACACGATGGCGCG 315  
DB 424 AGCACCAGGCGGACCAACGACACAGCGGACCAACAGACCAACGACCAAGC 483  
QY 316 GACAAATTTCTATCGCGCGCTGCGCAGCT--GACCAAGGCGCAAGCGCATCGCATC 372  
DB 484 AGCTGGAGAAAGATGCGGAGGCGAGATCAAGAACTGAGCTTCAACATCAACAC 543  
QY 373 GAGTTGACACGTCAATTCGACTTCCGCGCGGAGCTGAGAAAGCCTTACCGGCGTC 432  
DB 544 GTGCGCGACAGAGTGAAGAGATACAGCTGTCTCAACCTGAGCGTGTGCGCATC 603  
QY 433 GACTTGTGACATGACGAGCGCTGATGTGATGCGGACCATCAAGTGGCTGCAAGAG 492  
DB 604 GACACGACACGACGACGCTGATCCGCTGATCAATGCAACACAGCTGATCAACGAGCC 663  
QY 493 CAGAACTGATCCGCAAGGCGCGCGTGTGTGAGTGGCGCGCGCGCTGCGCGCT 552  
DB 664 TGCCCAAGGTGAGCTTCAGGCCCATCCCATTCATCACTGCGGCGCGCGCTTCGCG 723  
QY 553 GGCATCAAGGCGCGCTGCGGAGATGAAGTGGCATGCCACCAACCAATGCATGATC 612

DB 724 ATCTGAAGTGAAGACAAAGATTCAACGCGACCGGCGCTGCAAGAACGTGAGCAC 783  
QY 613 CCGGAGATGCCCAATCTCTCCCTCTGTGAGCTGATGAGACACTGACCTGTTCCAG 672  
DB 784 GTGCGTACACACCGAGCATCGCGCGCTGTGAGACACCGCTCTGTGAACGAGC 843  
QY 673 TCGGACATCAACACGAGCGGCGCAATCCGATCAACACCGCATGTGATTCGCGC 732  
DB 844 CTGGCGAGAGAGATGTGCTGCTGCTCGAGAACTTCACTGACACGCGCAAGACATC 903  
QY 733 GACATCTTTTGTGCTCAACACCTTCCGATGATCTTGTGCTATCAACCGCGCTGAGCGC 792  
DB 904 ATGTGACAGCTGAACGATCCGTGAGATCAACTCATTCATCCGCCCAACAAACGCGT 963  
QY 793 AGCTGTTCTGTGACCATGTGATGATGACGCGACCTGACATCTGGAGAAAGACTGGC 852  
DB 964 AAGACATTCACATGCGCGCGCGCGCTTCAACGCGCGGACATCATTCGCGGAC 1023  
QY 853 GTGCATGCGCGCGGCTGAGCTGATCAAGCGCGGCGCGCTGCAAGGACATCGGCATC 912  
DB 1024 ATCCGCGAGCGCCACTGCAACATGACAGGCAACGAGCACTGACCTGAGCAGATC 1083  
QY 913 GAGCTCAACGAGATGTACCGGAGTGGAGCTGTGATGATCCGCTTCCGCTATGAGC 972  
DB 1084 G---TGAGAACTGCGGAGACAGTTCGCAACAAAGACATCATCTTCAACAGCAGC 1140  
QY 973 CACTCTTGTGCGGTGTGTGCTGCACTACTACGCTGCGGAGCGCGGCTGAGCTGCGGAG 1032  
DB 1141 AGCGCGGCGGACCGCGAGATGTGTTCACAGCTTCACTGCGGCGGAGTTCCTTAC 1200  
QY 1033 GACATGACACCGAGCTGAACCGCGGATGTGTGCTCATGAGCCGATGTGATGCTG 1092  
DB 1201 TGCACACACGACGAGCTTTCACACGACCTGGAACATCAAGAGAGTGAACAAGACC 1260  
QY 1093 CCGAGAGGATGCCCCGATGCGGCGGCTATCGGAGACAGACATCTGATCGTGGAGAG 1152  
DB 1261 AAGAGAAAGACACCATCATCTCTGCGCTCGGACATCCGCGAGATCAACATGTGGCAG 1320  
QY 1153 GACGCTGCGGAGACATCAACGGGCTTCCGTTGCTCGGAAACACACCATTCGCA 1211  
DB 1321 GAGTGGGCAAGGCGCATGTACGCCCCCATTCGCGGCGACATCAAGTGACAGCA 1379

## RESULT 12

US-09-475-515-58  
Sequence 58, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 2181  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: gp140TM.modusa  
US-09-475-515-58

Query Match 5.0%; Score 60.6; DB 4; Length 2181;  
 Best Local Similarity 42.6%; Pred. No. 0.0025;  
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

196 CGCAAGTACGCGATGTCATCGACCAACAACCCGACGATTTGCGCGCGATGAC 255  
 Db CCCCTGTGCGTGAACCTGGAATGACCGACGACGACGACCAACGGAACCAAC 423  
 QY 256 GCGGCGACGCGCTGGCGCGGAGTTGGGCAACAATCACTACCGACTGGCGCGC 315  
 Db 424 AGCAACCGCGGACCAACAGACGACGACCAACAGCAACAGACGACGACGAC 483  
 QY 316 GACAATTTATGCGCGCGCGCGCGAGCT---GACCAAGCGCGCGCAAGCGCATCGCATC 372  
 Db 484 AGCTGGAGAAAGATGCCGAGGAGGAGATCAAGATGCAAGCTTCAACATCAACAGC 543  
 QY 373 GAGTTGACCAAGTCAATCTGACTTCCGCGCGCGCGCGCGCGCGCGCGCGCT 432  
 Db 544 GTGGCGGACAAAGGTGCAAGAGATACAGCTGTTCACAAAGTGGACGTGGGCCATC 603  
 QY 433 GACTTGTGACATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 492  
 Db 604 GACAACGCAACAGCGCACTACCGCTGATCACTGCAACACAGCGTATCAACCGCGC 663  
 QY 493 CAGAAGCTATCGCGGCAAGGCGCGCGGTGTGACGTGCGCGCGCGCGCGCGCT 552  
 Db 664 TGCCCCAAGGTGAGCTTGCAGCCCATCCATCCATCACTACGCGCGCGCGCGCT 723  
 QY 553 GCCATCAAGCT 612  
 Db 724 ATCTGGAAGTGAAGAGCAAGAGATTCAAGGCAACGCGCGCGCGCGCGCGCGCT 783  
 QY 613 CGCGAGATCGCGCAATGCTTCCCTTGTGTGAGCTGATGCAACCTGGACTGGTTCAG 672  
 Db 784 GTGCAAGTGAACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 843  
 QY 673 TCGGGGATCAACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 732  
 Db 844 CTGGCGGAGAGATGCTGCTGCGCTCGGAGATCTTCAACCAAGCGCGCGCGCT 903  
 QY 733 GACATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTTCAACCGCGCTGAGCGC 792  
 Db 904 ATGTGAGCTGAACGAGTCCGTGAGATCACTGCACTCGGCCCAACAAACAGCGCT 963  
 QY 793 ACGCTGTTTTCGACCATGTCATGACGCGACGCTCGACATCTGGGAGAAAGCTGGCC 852  
 Db 964 AAGAGCATCACTCGCGCGCGCGCGCGCGCGCTTCAACGCGCGCATCATCGCGAC 1023  
 QY 853 GTGCAATCGCT 912  
 Db 1024 ATCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1083  
 QY 913 GAGCTCAACGAGATGTACCCGCGAGTGGGACCTGCTGAAGTACCGCTTCCGCTATGGC 972  
 Db 1084 G---TGAAGAGCTGGCGCGAGTTCGGCAACAAGACATCATCTTCAACAGAGAGC 1140  
 QY 973 CACTCTTTCGCGCTGCTGCTGCTCACTACGCTGCGAGGCGCGCGTGGAGCTCGCGAG 1032  
 Db 1141 AGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTAC 1200  
 QY 1033 GACATGACACGAGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1092  
 Db 1201 TGCAACACGCGCGCGCTTCAACAGCACTGGAACATCAACGAGAGGAGCAAGACC 1260  
 QY 1093 CCGGAGGAGTGCCTGGTGCCTGCGCGCGCTATCGCGAGACGACATCTGATCGTGGGAG 1152  
 Db 1261 AAGAGAGAGCAACATCATCTGCGCTCGCGCATCGCGAGATCAACATGTGGCAG 1320  
 QY 1153 GAGGAGGAGAGCAACATCAACGCGCTTCCGCTGGTTCGGAAACACCAATCATCGGCA 1211  
 Db 1321 GAGGTGGGCGAGCATGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1379

RESULT 13  
 US-09-475-515-64  
 / Sequence 64, Application US/09475515A

/ Patent No. 6602705  
 / GENERAL INFORMATION:  
 / APPLICANT: BARNETT, Susan  
 / APPLICANT: ZUR MEGEDE, Jan  
 / APPLICANT: SRIVASTAVA, Indresh  
 / APPLICANT: LIAN, Ying  
 / APPLICANT: HARTOG, Karin  
 / APPLICANT: LIU, Hong  
 / APPLICANT: GREER, Catherine  
 / APPLICANT: SELBY, Mark  
 / APPLICANT: WALKER, Christopher  
 / TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
 / FILE REFERENCE: 1621.002  
 / CURRENT APPLICATION NUMBER: US/09/475,515A  
 / CURRENT FILING DATE: 1999-12-30  
 / NUMBER OF SEQ ID NOS: 90  
 / SOFTWARE: Patentln Ver. 2.0  
 / SEQ ID NO 64  
 / LENGTH: 2634  
 / TYPE: DNA  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: Description of Artificial Sequence: gp160.modUS4  
 US-09-475-515-64

Query Match 5.0%; Score 60.6; DB 4; Length 2634;  
 Best Local Similarity 42.6%; Pred. No. 0.0026;  
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

196 CGCAAGTACGCGATGTCATCGACCAACAACCCGACGATTTGCGCGCGATGAC 255  
 Db 364 CCCCTGTGCGTGAACCTGGAATGACCGACGACGACGACCAACGGAACCAAC 423  
 QY 256 GCGGCGACGCGCTGGCGCGGAGTTGGGCAACAATCACTACCGACTGGCGCGC 315  
 Db 424 AGCAACCGCGGACCAACAGACGACGACCAACAGCAACAGACGACGACGACGAC 483  
 QY 316 GACAATTTATGCGCGCGCGCGCGAGCT---GACCAAGCGCGCGCAAGCGCATCGCATC 372  
 Db 484 AGCTGGAGAAAGATGCCGAGGAGGAGATCAAGATGCAAGCTTCAACATCAACAGC 543  
 QY 373 GAGTTGACCAAGTCAATCTGACTTCCGCGCGCGCGCGCGCGCGCGCGCGCT 432  
 Db 544 GTGGCGGACAAAGGTGCAAGAGATACAGCTGTTCACAAAGTGGACGTGGGCCATC 603  
 QY 433 GACTTGTGACATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 492  
 Db 604 GACAACGCAACAGCGCACTACCGCTGATCACTGCAACACGCGGTATCAACCGCGC 663  
 QY 493 CAGAAGCTATCGCGGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 552  
 Db 664 TGCCCCAAGGTGAGCTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 723  
 QY 553 GCCATCAAGCT 612  
 Db 724 ATCTGGAAGTGAAGAGCAAGAGTTCAACGCGCACGCGCGCTGCAAGAGCTGAGACCC 783  
 QY 613 CGCGAGATGCGCAAAATCGTTCCTTTCGCGAGCTGATGAGAACCACTGAGACTGTTCCAG 672  
 Db 784 GTGCAAGTGAACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 843  
 QY 673 TCGGCGATCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 732  
 Db 844 CTGGCGGAGAGAGATGCTGCTGCTCGGAGAACTTCAACGAGAGCGCAAGACATC 903  
 QY 733 GACATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTTCAACCGCGCTGAGCGC 792  
 Db 904 ATGTGAGCTGAACGAGTCCGTGAGATCACTGACCGCGCGCGCGCGCGCGCGCT 963

QY 793 ACCTGTTCTGCGACCATGATGATGACCGCAGCTTCGACATCTGGGAGAAAGTGGCC 852  
DB 964 AAGAGCATCCACATCGGCCCCCGGCGCTTCTACGCCACCGGCAATCATCGCGAC 1023  
QY 853 GTGCATCGCGCGGGCTCGAGTGTATCAAGCCGGCGCGCGCTGCAGAGCATCGCCATC 912  
DB 1024 ATCCGCGAGGCCCACTGCAACATCAGCAAGGCCCACTGCAACACCTCGAGCATC 1083  
QY 913 GAGCTCAAGAGATGATACCGCGAGTGGGACCTCTGAAGTACCGCTCTTGGCTATGCG 972  
DB 1084 G---TGAGAGAGTGGCGCGAGCATTTGGCAACAAGCAATCATCTTCAACAGCAGC 1140  
QY 973 CACTCTTGGCGCTGCTGTGCTCACTAATGATGCGAGCGCGCGCTGAGCTGCGCGAG 1032  
DB 1141 AGCGGGGCGGACCCCGAGATCGTGTTCACAGCTTCACTGCGCGCGGAGTTCTTAC 1200  
QY 1033 GACATGACACCGAGTTGAAGCCCGGAGTGTCTCAATGAGCGATGTGATCTG 1092  
DB 1201 TGCACACAGCAGCGTGTTCACACAGCAGCTGGAATCATCAGAGAGGTGAACAAGACC 1260  
QY 1093 CCGGAGGCGATGCCCGGTGCGCGCGCTATCGGAGACAGCATCTGATCGTGGGAG 1152  
DB 1261 AAGGAGACGACACCATCATCTGCGCTGCGCGCATCGCGAGATCATCAATGTGGCAG 1320  
QY 1153 GACGTCGCGAGAACATCACCGGCTTCCGTTGCTCGGAGACACACATCATCCGCA 1211  
DB 1321 GAGGTGGGCAAGGCCATGTACGCCCCCGCATCGCGGCGCAATCAAGTGCAGACAA 1379

## RESULT 14

US-09-475-515-73  
; Sequence 73, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 73  
; LENGTH: 4766  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; ;  
; OTHER INFORMATION: gp160, modUS4, gag, modSF2  
; ;  
US-09-475-515-73

Query Match 5.0%; Score 60.6; DB 4; Length 4766;  
Best Local Similarity 42.6%; Pred. No. 0.00029;  
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CGCAAGTACGGCATGATCATGACCAACAACGCGCAGCATTTCCGCGCATTCAC 255  
DB 364 CCCCTGTGCTGACCTGAACTGACACCGACAGCTGACCGGACCAACGCGCAAC 423  
QY 256 GCGGCGCAGCCTTGGCGCGGAGCTTGGCGCAACATTAACCGCATGCGGCGGC 315  
DB 424 AGCAGCAAGCGGACCAACAGCACGCGGCGCAACAGACCAAGACCAAGCAGCAG 483  
QY 316 GACAAATTTCTATCGCGCGTGGCGAGCT---GACCAAGGCGGCGCAAGCGCATCGGCATC 372

DB 484 AGCTGGAGAGAGTGGCCGAGGCGGAGATCAAGAACTGAGGTTCAACATCACACAGC 543  
QY 373 GAGTTGACACAGTCAATTCGACTTCCGCGCGAGCTTGAGAGACCTTACCGGCGCTC 432  
DB 544 GTGCGGACAGAGTGCAGAGAGATGACAGCTGTTCTACAGCTGAGCTGTGCGCATC 603  
QY 433 GACTTGTGAGATGACGCGAGCCCTGATGTGGAATGCGGACCATCAAGTGGCTGGAAG 492  
DB 604 GACAAACACAGCAGCTTACCGCTGATCAACTGACACACAGCTGATCAACAGGCC 663  
QY 493 CAGAACTGATTCGCGAAGGCGCGCGGTGTGTGACGTTCGCGCGCGCGCTGCGGCGCT 552  
DB 664 TCGCCCAAGGTGAGGTGAGGCTTACGCCATCCCATCACTATGAGGCCCGCGCTTGGCC 723  
QY 553 GCCATCAAGGCGCGGTGCGCGAGCATGAATGCGGATGCCACCAATGCGATGATC 612  
DB 724 ATCTGAAGTGCAGAGCAAGAAATTCACCGCACCGGCGCTTGCAGAAACGTGACACC 783  
QY 613 CGCGAGATCGCAAAATCGTTCCCTTGTGAGCTGATGGAACACTGACCTGGACCTGGTTCAAG 672  
DB 784 GTGCAATGACACCAAGGATCGCGCGGTGTGAGACACCACTGCTGTGAACGCGAC 843  
QY 673 TCGGCGATCAACACGAGCGCGGCAATCCGTTACCAACCGCATCTGATTCGCGG 732  
DB 844 CTGGCGAGAGAGATGATGTGTGCTGCGCTCCGAGAACTTCAACGACAAAGCCAAAGCATC 903  
QY 733 GACATCTTTGCTTAACCTTCCGATGATCTTGGCTACTACCGCGCTGAGCGC 792  
DB 904 ATCTGACAGCTGACAGATCGCTGAGATCACTGACATCCGCCCAACAAACAGCGCT 963  
QY 793 AGCTGTTCTGCGACATGATGATGACGCGCGCTTCGATCTGAGAGAAAGTGGCC 852  
DB 964 AAGAGCATCAATGATGCGCGCGCGCTTCTACGCGACCGGCAATCATCGCGAC 1023  
QY 853 GTGCATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAGGACATTCGCATC 912  
DB 1024 ATCCGCGAGGCGCACTGCAATCAAGAGGCAATGGAACCAACCTCGAGCAGATC 1083  
QY 913 GAGCTCAAGAGATGATACCGGAGTGGGACCTGCTGAAGTACCGCTCTTGGCTATGCG 972  
DB 1084 G---TGAGAGAGCTGCGGAGCAGATTCGCGCAACAAGACCATCATCTTCAACAGCAGC 1140  
QY 973 CACTCTTGGCGGTGCTGTGCACTACAGGTGCGGCGCGGCGGTGAGAGCTGCGCGAG 1032  
DB 1141 AGCGCGCGGACCCCGAGATGTGTTCCACAGCTTCACTGCGCGCGGAGTTCTTAC 1200  
QY 1033 GACATGACACCGAGCTGAAGCCCGGCGATGTGTCTCATGAGCCGATGTGATGCTG 1092  
DB 1201 TGCACACAGCAGCTGTTCACAGCAGCTGGAACATCACGAGAGGTGAACAAGACC 1260  
QY 1093 CCGAGGCGATGCCCGGTGCGCGGCTATGCGGACGACATCTGATGCTCGGAGAG 1152  
DB 1261 AAGGAGACGACACCATCATCTGCGCTGCGCGCATCCGAGATCATCAATGTGGCAG 1320  
QY 1153 GACGTCGCGAGAACATCACCGGCTTCCGTTGCTCGGAGACACACATCATTCGCA 1211  
DB 1321 GAGGTGGGCAAGGCCATGTACGCCCCCGCATCGCGGCGCAATCAAGTGCAGACAA 1379

## RESULT 15

US-09-025-691-4  
; Sequence 4, Application US/09025691  
; Patent No. 6069299  
; GENERAL INFORMATION:  
; APPLICANT: Broadway, Roxanne M.  
; APPLICANT: Harman, Gary E.  
; TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH  
; TITLE OF INVENTION: CHITINOLYTIC ENZYMES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using 6w model

Run on: November 19, 2004, 12:40:44 ; Search time 673 Seconds  
(without alignments)  
9731.339 Million cell updates/sec

Title: US-10-807-228-2

Perfect score: 1212  
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACCATCATCGGCAAC 1212

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	981.6	81.0	1215	US-10-251-078-1	Sequence 1, Appl1
2	981.6	81.0	1215	US-10-251-078-17	Sequence 17, Appl1
3	978.4	80.7	1215	US-10-251-078-9	Sequence 9, Appl1
4	976.8	80.6	1212	US-10-251-078-11	Sequence 11, Appl1
5	976.8	80.6	1212	US-10-251-078-13	Sequence 13, Appl1
6	976.8	80.6	1215	US-10-251-078-15	Sequence 15, Appl1
7	975.2	80.5	1212	US-10-251-078-21	Sequence 21, Appl1
8	973.6	80.3	1212	US-10-251-078-19	Sequence 19, Appl1
9	75.4	6.2	2733	US-10-437-963-88003	Sequence 88003, A
10	73.4	6.1	1262	US-10-425-115-109294	Sequence 109294, A
11	69.8	5.8	1483	US-10-437-963-38426	Sequence 38426, A
12	69.6	5.7	9025608	US-10-156-761-1	Sequence 1, Appl1

13	68.8	5.7	65140	17	US-10-203-295-1	Sequence 1, Appl1
14	68.8	5.7	125401	17	US-10-203-295-35	Sequence 35, Appl1
15	68.6	5.7	1169	17	US-10-437-963-57011	Sequence 57011, A
16	68	5.6	834	17	US-10-156-761-4932	Sequence 4932, Ap
17	68	5.6	1212	18	US-10-411-910A-222	Sequence 222, Ap
18	67	5.5	2569	17	US-10-437-963-88004	Sequence 88004, A
19	67	5.5	4512	15	US-10-156-761-2222	Sequence 2222, Ap
20	67	5.5	9025608	15	US-10-156-761-1	Sequence 1, Appl1
21	66.8	5.5	2712	16	US-10-282-122A-24663	Sequence 24663, A
22	66	5.4	2037	15	US-10-228-063-48	Sequence 48, Appl1
23	65	5.4	573	15	US-10-156-761-3931	Sequence 3931, Ap
24	65	5.4	2026	17	US-10-437-963-32278	Sequence 32278, A
25	65	5.4	2052	17	US-10-282-122A-11731	Sequence 11731, A
26	65	5.4	2649	17	US-10-437-963-112	Sequence 112, App
27	64.6	5.3	1752	18	US-10-411-910A-234	Sequence 234, App
28	64.4	5.3	2960	17	US-10-437-963-34365	Sequence 34365, A
29	64	5.3	1516	17	US-10-437-963-29365	Sequence 29365, A
30	64	5.3	2364	17	US-10-473-687-3	Sequence 3, Appl1
31	63.6	5.2	690	13	US-10-006-922-27	Sequence 27, Appl1
32	63.6	5.2	690	14	US-10-081-864-19	Sequence 19, Appl1
33	63.6	5.2	1248	18	US-10-411-910A-220	Sequence 28, App
34	63.6	5.2	1311	18	US-10-411-910A-200	Sequence 20, App
35	63.6	5.2	9546	16	US-10-463-074-2	Sequence 2, Appl1
36	63.4	5.2	1083	16	US-10-260-238-271	Sequence 271, App
37	63	5.2	18435	15	US-10-156-761-412	Sequence 412, App
38	63	5.2	100000	15	US-10-156-761-15103	Sequence 15103, A
39	62.8	5.2	1053	15	US-10-156-761-5241	Sequence 5241, Ap
40	62.8	5.2	9785	16	US-10-359-120-176	Sequence 176, App
41	62.6	5.2	1377	18	US-10-411-910A-266	Sequence 266, App
42	62.6	5.2	1479	15	US-10-156-761-4461	Sequence 4461, App
43	62.6	5.2	1758	18	US-10-411-910A-209	Sequence 209, App
44	62.4	5.1	1294	9	US-09-748-033-2	Sequence 2, Appl1
45	62.4	5.1	1335	15	US-10-156-761-7186	Sequence 7186, Ap

ALIGNMENTS

RESULT 1

US-10-251-078-1

Sequence 1, Application US/10251078

Publication No. US20030119084A1

GENERAL INFORMATION:

APPLICANT: Roche Diagnostics GmbH

APPLICANT: Shao, Zhixin

APPLICANT: Schmuck, Rainer

APPLICANT: Kratzsch, Peter

APPLICANT: Kenkies, Janet

APPLICANT: Weisser, Harald

TITLE OF INVENTION: Variants of an Erwina-type creatinase

FILE REFERENCE: 20981 EP

CURRENT FILING DATE: 2002-09-20

CURRENT APPLICATION NUMBER: US/10/251.078

NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1215

TYPE: DNA

ORGANISM: Erwina sp.

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1215)

US-10-251-078-1

Query Match

81.0% Score 981.6; DB 15; Length 1215;

Best Local Similarity 88.1% Pred. No. 8.9e-254;

Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy

1 ATGACTGACGACATGTTGCAAGATGACACACGCGGAGAAAGATTATTCGCG 60

|||||

Db

1 ATGACTGACGACATGTTGCAAGATGACACACATGCTGAGAGGAATATTCGCC 60

|||||

Qy

61 TTTTCGATGCGGAGATGACCCGCCCAAAACGACGTTCCGCGCTGATGCGCAAGAC 120

Db 61 TTTTCGATGCGGAGATGACGGCCCCGAGTGAATGGGGCTGGATGGCCGAAAC 120  
Qy 121 AATGTCGATGCGGCGCTGTTCACCTTATCACTGATCACTATTCGGCTGGCTG 180  
Db 121 GAGTCGACGCTGGCTGTTCACCTTATCACTGATCACTATTCGGATTCCTG 180  
Qy 181 TACTGCTATTTGGAGCGCAAGTACGGCATGCTCATATGACCAACAAGCCACAGCAT 240  
Db 181 TACTGCTATTTGGAGCGCAAGTACGGCATGCTCATATGACCAACAAGCCACAGCAT 240  
Qy 241 TGGGCGGCGATGACCGGCGCGAGCCCTGAGCGCGGAGCTTGGCGGCAACATCACTAC 300  
Db 241 TGGGCGGCGATGACCGGCGCGAGCCCTGAGCGCGGAGCTTGGCGGCAACATCACTAC 300  
Qy 301 ACCGACTGGCGCGCGACAAATTTATCGCGCGGTGGCGCAAGTGAACAAGCGCGCAAG 360  
Db 301 ACCGACTGGCGCGCGCGACAAATTTATCGCGCGGTGGCGCAAGTGAACAAGCGCGCAAG 360  
Qy 361 CGCATCGGATCGAGTTGCAACACGTCATCTCGATTTCCGCGCGCAAGTGAAGGCC 420  
Db 361 CGCATCGGATCGAGTTGCAACACGTCATCTCGATTTCCGCGCGCAAGTGAAGGCC 420  
Qy 421 CTACCGGCGGCTGCACTTCGTGACACGCGCTCGATGAGTGGCAACATCAAG 480  
Db 421 CTGCGCGGCGGCTGCACTTCGTGATCGATCGATCGATGATGCGCAACGCTCAAG 480  
Qy 481 TCGCTCGAAGACAGAAAGTATCGCGAAGCGCGCGGTGTGTGACGTGGCGCGCG 540  
Db 481 TCGCTCGAAGACAGAAAGTATCGCGAAGCGCGCGGTGTGTGACGTGGCGCGGTGCC 540  
Qy 541 GCGTCGCGGCGCTGCTCATCAAGCGCGCGGTGGCGAGCAAGTGGCGATCGCCACAC 600  
Db 541 GCGTCGCGGCGCTGCTCATCAAGCGCGCGGTGGCGAGCAAGTGGCGATCGCCACAC 600  
Qy 601 AATGCGATGATCGCGAGATCGCAAAATGTTCCCTTGTGTGAGTGTATGACACTGG 660  
Db 601 AATGCGATGATCGCGAGATCGCAAAATGTTCCCTTGTGTGAGTGTATGACACTGG 660  
Qy 661 ACCTGATTCAGTCCGGGATCAACACGCGCGCGCAATCGGTCAACAACGCGCATC 720  
Db 661 ACCTGATTCAGTCCGGGATCAACACGCGCGCGCAATCGGTCAACAACGCGCATC 720  
Qy 721 GTGCAATCGGCGGATCTTCTTGTGCTCAACACCTTCCGATGATCTTGGCTACTAC 780  
Db 721 GTGCAATCGGCGGATCTTCTTGTGCTCAACACCTTCCGATGATCTTGGCTACTAC 780  
Qy 781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGAGCGGAGCGCTGCACTGGAG 840  
Db 781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGAGCGGAGCGCTGCACTGGAG 840  
Qy 841 AAGAACTGGCGCGGATCGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900  
Db 841 AAGAACTGGCGCGGATCGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900  
Qy 901 GACATGCGCATGAGTCAACGAGATGACCGGAGTGGGACCTGCTGAAGTACCGCTCC 960  
Db 901 GATATGCGCATGAGTCAACGAGATGACCGGAGTGGGATGCTGCTGAAGTACCGCTCC 960  
Qy 961 TTGGGCTATGGGCACTCCCTTGGCGGCTGTGCGCACTAAGTGGCGGAGCGCGCG 1020  
Db 961 TTGGGCTATGGGCACTCCCTTGGCGGCTGTGCGCACTAAGTGGCGGAGCGCGCG 1020  
Qy 1021 GAGTCGCGGAGACATCGACACCGAGCTGAAGCCCGGATGGTCTCAATGAGCGCG 1080  
Db 1021 GAGTCGCGGAGACATCGACACCGAGCTGAAGCCCGGATGGTCTCAATGAGCGCG 1080  
Qy 1081 AATGTCATGCTGCGGAGGCGATGCGCGGTGGCGGCTTATGCGAGACATCACTCTG 1140  
Db 1081 AATGTCATGCTGCGGAGGCGATGCGCGGTGGCGGCTTATGCGAGACATCACTCTG 1140  
Qy 1141 ATGCTCGGAGGAGAGGTGCGGAGACATCAACGCGCTTCCGTTGGGTGCGGACACACC 1200  
Db 1141 ATGCTCGGAGGAGAGGTGCGGAGACATCAACGCGCTTCCGTTGGGTGCGGAGACACACC 1200

Db 1141 ATGCTGGGAGGAGAGGCGCGGAGAGACATTAACGGATTCCTTGGGCGCTGAGACAAAC 1200  
Qy 1201 ATCATCCGCAAC 1212  
Db 1201 ATCATCCGCAAC 1212  
RESULT 2  
US-10-251-078-17  
; Sequence 17, Application US/10251078  
; Publication No. US20030119084A1  
; GENERAL INFORMATION:  
; APPLICANT: Roche Diagnostics GmbH  
; APPLICANT: Shao, Zhixin  
; APPLICANT: Schmuck, Rainer  
; APPLICANT: Kraltsch, Peter  
; APPLICANT: Kenkies, Janet  
; APPLICANT: Weisner, Harald  
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase  
; FILE REFERENCE: 20981 BP  
; CURRENT APPLICATION NUMBER: US/10/251,078  
; CURRENT FILING DATE: 2002-09-20  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 17  
; LENGTH: 1215  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:variant Ctrc2  
; NAME/KEY: CDS  
; LOCATION: (1)..(1215)  
US-10-251-078-17  
Query Match 81.0%; Score 981.6; DB 15; Length 1215;  
Best Local Similarity 88.1%; Pred. No. 8.9e-254;  
Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;  
Qy 1 ATGACTGACGACATGTTGACAGTGAATGACCAACGCGGAGAAAGATTTTCGCCG 60  
Db 1 ATGACTGACGACATGTTGACAGTGAATGACCAACATGATGAGAAATATTCGCC 60  
Qy 61 TTTTTCGATGCGGAGATACCCCGCGCAAAACAGCTTTCGCGGTGGATGGCCAAAG 120  
Db 61 TTTTTCGATGCGGAGATACCCCGCGCAAAACAGCTTTCGCGGTGGATGGCCAAAG 120  
Qy 121 AATGCGATGCGGCGGCTGTTCACCTCTATACATGATCACTAATTTCCGCTGGCTG 180  
Db 121 GAGTCGAGCGCTGCGCTGTTCACCTCTATACATGATCACTAATTTCCGCTGGCTG 180  
Qy 181 TACTGCTATTTGAGCGCAAGTACGGCATGCTCATATGACCAACAAGCCACAGCAT 240  
Db 181 TACTGCTATTTGAGCGCAAGTACGGCATGCTCATATGACCAACAAGCCACAGCAT 240  
Qy 241 TGGGCGGCGATGACCGGCGCGAGCCCTGAGCGCGGAGCTTGGCGGCAACATCACTAC 300  
Db 241 TGGGCGGCGATGACCGGCGCGAGCCCTGAGCGCGGAGCTTGGCGGCAACATCACTAC 300  
Qy 301 ACCGACTGGCGCGCGGACAAATTTATCGCGCGGTGGCGCAAGTGAACAAGCGCGCAAG 360  
Db 301 ACCGACTGGCGCGCGGACAAATTTATCGCGCGGTGGCGCAAGTGAACAAGCGCGCAAG 360  
Qy 361 CGCATCGGATCGAGTTGCAACACGTCATCTCGATTTCCGCGCGCAAGTGAAGGCC 420  
Db 361 CGCATCGGATCGAGTTGCAACACGTCATCTCGATTTCCGCGCGCAAGTGAAGGCC 420  
Qy 421 CTACCGGCGGCTGCACTTCGTGACACGCGCTCGATGAGTGGCAACATCAAG 480  
Db 421 CTGCGCGGCGGCTGCACTTCGTGATCGATCGATGATGCGCAACGCTCAAG 480  
Qy 481 TCGCTCGAAGACAGAAAGTATCGCGAAGCGCGGTGTGTGACGTGGCGCGCG 540  
Db 481 TCGCTCGAAGACAGAAAGTATCGCGAAGCGCGGTGTGTGACGTGGCGCGCG 540



```

Db      481 TCGCTCGAAGAGCAGAGGTGATCCGCGAGGGTGCCTGATCTGCGACGTGGCGGTGC 540
Qy      541 GCGTGGCGGCTGCTCAAGGCGGCGTCCCGACGATGAGAGTGGCGATCGCACACC 600
Db      541 GCGTGGCGGCTGCTCAAGGCGGCGTCCCGACGATGAGAGTGGCGATCGCACACC 600
Qy      601 AATGCGATGATCCGCGAGATCCGCAATGTTCCCTTGGTGGAGTGAATGACACTGG 660
Db      601 AATGCGATGATCCGCGAGATCCGCAATGTTCCCTTGGTGGAGTGAATGACACTGG 660
Qy      661 ACCTGGTTCAGTCCGGGATCAACACCGA CCGCGCGACAAATCCGGTCA CCAACCGCATC 720
Db      661 ACCTGGTTCAGTCCGGGATCAACACCGA CCGCGCGACAAATCCGGTCA CCAACCGCATC 720
Qy      721 GTGCAATCCGGGAGCATCTTTTCGCTCAACACTTCCGATGATCTTGGCTA CTAACCC 780
Db      721 GTGCAATCCGGGAGCATCTTTTCGCTCAACACTTCCGATGATCTTGGCTA CTAACCC 780
Qy      781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
Db      781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
Qy      841 AAGAACGTGCGCGTGCATGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Db      841 AAGAACGTGCGCGTGCATGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Qy      901 GACATCGCATGACATCTCAACGAGATGTAACCGCGAGTGGAGACTCTGTAAGTACCGCTCC 960
Db      901 GATATCGCATGACATCTCAACGAGATGTAACCGCGAGTGGAGACTCTGTAAGTACCGCTCC 960
Qy      961 TTCGCTATGCGCATCTCTTCCGCGTGTGCTGCTGCTCACTACTACGTCGCGAGCGCGCTG 1020
Db      961 TTCGCTATGCGCATCTCTTCCGCGTGTGCTGCTGCTCACTACTACGTCGCGAGCGCGCTG 1020
Qy      1021 GAGCTGCGGAGAGCATGACACCGAGCTGAAGCCCGCATGCTGCTCACTGAGAGCG 1080
Db      1021 GAACTGCGGAGAGCATGATACGCTGCTGACAGCCCGCATGCTGCTCACTGAGAGCG 1080
Qy      1081 ATGGGATGCTGCGCGAGGCGATGCGCGTGCCTGCGCGCTATCCGAGACAGCATCTCTG 1140
Db      1081 ATGGGATGCTGCGCGAGGCGATGCGCGTGCCTGCGCGCTATCCGAGACAGCATCTCTG 1140
Qy      1141 ATCGTCGGGAGAGCGTGCAGCAACACCGCTTCCGCTGCTGCGGACACACAC 1200
Db      1141 ATCGTCGGGAGAGCGTGCAGCAACACCGCTTCCGCTGCTGCGGACACACAC 1200
Qy      1201 ATCATCCGCAAC 1212
Db      1201 ATCATCCGCAAC 1212

```

```

RESULT 3
US-10-251-078-9
; Sequence 9, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schumuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251.078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:

```

```

; OTHER INFORMATION: Description of Artificial Sequence: variant CT1m24
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1215)
US-10-251-078-9

```

```

Query Match      80.7%; Score 978.4; DB 15; Length 1215;
Best Local Similarity 88.0%; Pred. No. 6,5e-253;
Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

```

```

Qy      1 ATGACTGACGACATTTTCAGTGAATAATGACCAACGCGGAGAAAGATATTTCGCG 60
Db      1 ATGACTGACGACATTTTCAGTGAATAATGACCAACGCGGAGAAAGATATTTCGCG 60
Qy      61 TTTTCGATGCCGAGATGACCGCGCGCAAAACGAGTTCCGGGCTGAGTGGCAAGAAC 120
Db      61 TTTTCGATGCCGAGATGACCGCGCGCAAAACGAGTTCCGGGCTGAGTGGCAAGAAC 120
Qy      121 AATGTCGATGGCGGCTTTTCACTTATCACTGACATCACTACTATTTCGGCTGG 180
Db      121 AATGTCGATGGCGGCTTTTCACTTATCACTGACATCACTACTATTTCGGCTGG 180
Qy      181 TACTGCTATTTTCGAGCGAAGTACGGCATGATGACACCAACGAGCCAGCAT 240
Db      181 TACTGCTATTTTCGAGCGAAGTACGGCATGATGACACCAACGAGCCAGCAT 240
Qy      241 TCGGCGGCGATGACGCGGCGCAAGCCTGCGCGCGACGCTTGGCGCAACATCACTAC 300
Db      241 TCGGCGGCGATGACGCGGCGCAAGCCTGCGCGCGACGCTTGGCGCAACATCACTAC 300
Qy      301 ACCGATCGGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCATGACCAACGCGCGCG 360
Db      301 ACCGATCGGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCATGACCAACGCGCGCG 360
Qy      361 CGCATCGGATGATTCGACGATTCGACGATTCGACGATTCGACGATTCGACGATTCG 420
Db      361 CGCATCGGATGATTCGACGATTCGACGATTCGACGATTCGACGATTCGACGATTCG 420
Qy      421 CTACCGGCGCTGACCTTCTGATGACATGACGACGCTTGTGATGATGATGATGATG 480
Db      421 CTACCGGCGCTGACCTTCTGATGACATGACGACGCTTGTGATGATGATGATGATG 480
Qy      481 TCGCTCGAAGAGCAGAGCTGATCCGCGAGGCGCGCGCGCTGATGATGATGATGATG 540
Db      481 TCGCTCGAAGAGCAGAGCTGATCCGCGAGGCGCGCGCGCTGATGATGATGATGATG 540
Qy      541 GCGTGGCGGCTGCTCAAGGCGGCGTCCCGAGCATGAGAGTGGCGATCGCACACC 600
Db      541 GCGTGGCGGCTGCTCAAGGCGGCGTCCCGAGCATGAGAGTGGCGATCGCACACC 600
Qy      601 AATGCGATGATCCGCGAGATCCGCAATGTTCCCTTGGTGGAGTGAATGACACTGG 660
Db      601 AATGCGATGATCCGCGAGATCCGCAATGTTCCCTTGGTGGAGTGAATGACACTGG 660
Qy      661 ACCTGGTTCAGTCCGGGATCAACACCGA CCGCGCGACAAATCCGGTCA CCAACCGCATC 720
Db      661 ACCTGGTTCAGTCCGGGATCAACACCGA CCGCGCGACAAATCCGGTCA CCAACCGCATC 720
Qy      721 GTGCAATCCGGGAGCATCTTTTCGCTCAACACTTCCGATGATCTTGGCTA CTAACCC 780
Db      721 GTGCAATCCGGGAGCATCTTTTCGCTCAACACTTCCGATGATCTTGGCTA CTAACCC 780
Qy      781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
Db      781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
Qy      841 AAGAACGTGCGCGTGCATGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Db      841 AAGAACGTGCGCGTGCATGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Qy      901 GACATCGCATGACATCTCAACGAGATGTAACCGCGAGTGGAGACTCTGTAAGTACCGCTCC 960
Db      901 GATATCGCATGACATCTCAACGAGATGTAACCGCGAGTGGAGACTCTGTAAGTACCGCTCC 960

```

QY 961 TTGGCTATGCGCACTCTTCCGGCGTGTGCTGCGCACTACACGCTGCGGAGCGCGCTG 1020  
DB 961 TTGGCTATGCGCACTCTTCCGGCGTGTGCTGCGCACTACACGCTGCGGAGCGCGCTG 1020  
QY 1021 GAGCTCGGAGGACATCGACACCGAGCTGAAGCCCGGCAATGCTGCTTCATGAGCGG 1080  
DB 1021 GAACTCGGAGGACATCGACATCCGCTGCTGCGAGCCCGGCAATGCTGCTTCATGAGCGG 1080  
QY 1081 ATGCTATGCTGCTGCGGAGGCAATCCCGCTGCGGCGCTGCTGCGAGACGACATCTTG 1140  
DB 1081 ATGCTATGCTGCTGCGGAGGCAATCCCGCTGCGGCGCTGCTGCGAGACGACATCTTG 1140  
QY 1141 ATGCTGCGGAGGAGCGGCTGCGGAGACATCGACCGGCTGCTGCTGCGGAGACACAC 1200  
DB 1141 ATGCTGCGGAGGAGCGGCTGCGGAGACATCGACCGGCTGCTGCTGCGGAGACACAC 1200  
QY 1201 ATCATCGGCAC 1212  
DB 1201 ATCATCGGCAC 1212

RESULT 4  
US-10-251-078-11  
; Sequence 11, Application US/10251078  
; Publication No. US20030119084A1  
; GENERAL INFORMATION:  
; APPLICANT: Roche Diagnostics GmbH  
; APPLICANT: Shao, Zhixin  
; APPLICANT: Schuck, Rainer  
; APPLICANT: Kratzsch, Peter  
; APPLICANT: Kenkies, Janet  
; APPLICANT: Weisser, Harald  
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase  
; FILE REFERENCE: 20981 EP  
; CURRENT APPLICATION NUMBER: US/10/251,078  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 1212  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m9  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1212)  
US-10-251-078-11

Query Match 80.6%; Score 976.8; DB 15; Length 1212;  
Beet Local Similarity 87.9%; Pred. No. 1.7e-252;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGATGACGACATGTTGACGTGATGAATGACCAACGCGAGAAATATTTCGCGG 60  
DB 1 ATGATGACGACATGTTGACGTGATGAATGACCAACGCGAGAAATATTTCGCGG 60  
QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGCGCAAGAC 120  
DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGCGCAAGAC 120  
QY 121 AATGTCGATGCGGCGCTGTTGACCTTATCACTGTCATCACTATCTTCGCGCTG 180  
DB 121 AATGTCGATGCGGCGCTGTTGACCTTATCACTGTCATCACTATCTTCGCGCTG 180  
QY 121 GACGTGAGCGCTGCTGTTGACCTTATCACTGTCATCACTATCTTCGCGCTG 180  
DB 121 GACGTGAGCGCTGCTGTTGACCTTATCACTGTCATCACTATCTTCGCGCTG 180  
QY 181 TACTGCTATTTGCGAGCGAATGACGATGTCATGACCAACAGCGACGAGATT 240  
DB 181 TACTGCTATTTGCGAGCGAATGACGATGTCATGACCAACAGCGACGAGATT 240  
QY 241 TCGGCGGCGATGACG 300  
DB 241 TCGGCGGCGATGACG 300

QY 301 ACCGACTGCGCGCGGACCAATTTCTATGCGCGCGCTGCGCGCACTGACCAACGCGCGCGG 360  
DB 301 ACCGACTGCGCGCGGACCAATTTCTATGCGCGCGCTGCGCGCACTGACCAACGCGCGCGG 360  
QY 361 CGCATCGGATGAGTTCGACCAAGCTGATCTGCACTTCGCGCGCGCGCGCGCGCGCGCG 420  
DB 361 CGCATCGGATGAGTTCGACCAAGCTGATCTGCACTTCGCGCGCGCGCGCGCGCGCGCG 420  
QY 421 CTACCGGCGCTGCACTTCGTCGACATGACCGCGCGCTGATGATGATGATGATGATGAT 480  
DB 421 CTACCGGCGCTGCACTTCGTCGACATGACCGCGCGCTGATGATGATGATGATGATGAT 480  
QY 481 TCGCTCGAAGAGCAAACTGATTCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB 481 TCGCTCGAAGAGCAAACTGATTCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
QY 541 GCGTGGCGGCTGCGCATGAGCGCGCGCTGCGCGAGATGAGAGTGGCGATCGCAACAC 600  
DB 541 GCGTGGCGGCTGCGCATGAGCGCGCGCTGCGCGAGATGAGAGTGGCGATCGCAACAC 600  
QY 601 AATGCGATGATTCGCGGAGATGCGCAATGCTTCCCTTCTGAGAGCTGATGAGACACTGG 660  
DB 601 AATGCGATGATTCGCGGAGATGCGCAATGCTTCCCTTCTGAGAGCTGATGAGACACTGG 660  
QY 661 ACCGTGTTCCAGTGGCGGATGATCAACCGAGCGCGCGCAATCGCGTCAACCAACCGCATC 720  
DB 661 ACCGTGTTCCAGTGGCGGATGATCAACCGAGCGCGCGCAATCGCGTCAACCAACCGCATC 720  
QY 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTCGCGAGATGATTCGCGCTACTACAC 780  
DB 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTCGCGAGATGATTCGCGCTACTACAC 780  
QY 781 GCGCTGAGCGGACGCTGCTGCGGACATGATGAGAGCGCGCGCGCGCGCGCGCGCGCG 840  
DB 781 GCGCTGAGCGGACGCTGCTGCGGACATGATGAGAGCGCGCGCGCGCGCGCGCGCGCG 840  
QY 841 AAGAACGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 AAGAACGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 GACATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
DB 901 GACATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
QY 961 TTTGCGTATGCGGACACTCTTTCGCGGCTGCTGCGCACTATGAGTGGAGTGGAGTGG 1020  
DB 961 TTTGCGTATGCGGACACTCTTTCGCGGCTGCTGCGCACTATGAGTGGAGTGGAGTGG 1020  
QY 1021 GAGCTGCGGAGGACATGACACGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 1021 GAGCTGCGGAGGACATGACACGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
QY 1081 ATGCTGATGCTGCGGAGGAGGATGCGCGGCTGCGCGGCTGATGCGAGACGACATCTTG 1140  
DB 1081 ATGCTGATGCTGCGGAGGAGGATGCGCGGCTGCGCGGCTGATGCGAGACGACATCTTG 1140  
QY 1141 ATGCTGCGGAGGAGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATG 1200  
DB 1141 ATGCTGCGGAGGAGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATG 1200  
QY 1201 ATCATCGGCAC 1212  
DB 1201 ATCATCGGCAC 1212

RESULT 5  
US-10-251-078-13  
; Sequence 13, Application US/10251078  
; Publication No. US20030119084A1  
; GENERAL INFORMATION:  
; APPLICANT: Roche Diagnostics GmbH  
; APPLICANT: Shao, Zhixin

APPLICANT: Schmuck, Rainer  
APPLICANT: Kratzsch, Peter  
APPLICANT: Kenkies, Janet  
APPLICANT: Weisner, Harald  
TITLE OF INVENTION: Variants of an Erwinia-type creatinase  
FILE REFERENCE: 20981 EP  
CURRENT APPLICATION NUMBER: US/10/251,078  
CURRENT FILING DATE: 2002-09-20  
NUMBER OF SEQ. ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ. ID NO. 13  
LENGTH: 1212  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: variant CT2m10  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1212)  
US-10-251-078-13

Query Match 80.6%; Score 976.8; DB 15; Length 1212;  
Best Local Similarity 87.9%; Pred. No. 1.7e-252;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGAAATGCAACAACGCGAGAAAGATTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGAAATGCAACAATGCGAAGAAATATTCGCG 60  
QY 61 TTTTCGGATCGGAGATATACCGCGGCAAGCGTTGCGGCGGATGCGCAAGAAC 120  
DB 61 TTTTCGGATCGGAGATATACCGCGGCAAGCGTTGCGGCGGATGCGCAAGAAC 120  
QY 121 AATGTCGACGCGCGCTGTTCACTTATCACTGATCACTATTCGCGTGGCTG 180  
DB 121 GACGTGACGCGCGCTGTTCACTTATCACTGATCACTATTCGCGTGGCTG 180  
QY 121 GACGTGACGCGCGCTGTTCACTTATCACTGATCACTATTCGCGTGGCTG 180  
DB 121 GACGTGACGCGCGCTGTTCACTTATCACTGATCACTATTCGCGTGGCTG 180  
QY 181 TACTGCTATTTGCGACGCAATGCGGATGCTGATGACCAACACCGCCGACGATT 240  
DB 181 TACTGCTATTTGCGACGCAATGCGGATGCTGATGACCAACACCGCCGACGATT 240  
QY 241 TCGCGCGGATGACG 300  
DB 241 TCGCGCGGATGACG 300  
QY 241 TCGCGCGGATGACG 300  
DB 241 TCGCGCGGATGACG 300  
QY 301 ACCGATCGGCG 360  
DB 301 ACCGATCGGCG 360  
QY 301 ACCGATCGGCG 360  
DB 301 ACCGATCGGCG 360  
QY 361 CGCATCGGATGACG 420  
DB 361 CGCATCGGATGACG 420  
QY 421 CTACCGGCG 480  
DB 421 CTACCGGCG 480  
QY 421 CTACCGGCG 480  
DB 421 CTACCGGCG 480  
QY 481 TCGCTCGAAGACAGAGCTGATCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB 481 TCGCTCGAAGACAGAGCTGATCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 540  
QY 481 TCGCTCGAAGACAGAGCTGATCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB 481 TCGCTCGAAGACAGAGCTGATCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 540  
QY 541 GCGTCG 600  
DB 541 GCGTCG 600  
QY 601 AATGCGATGATCGCGGAGATGCGCAATGTTCCCTTGTGAGCTGATGACCTGG 660  
DB 601 AATGCGATGATCGCGGAGATGCGCAATGTTCCCTTGTGAGCTGATGACCTGG 660  
QY 661 ACCTGTTCCAGTCCGGGACATCAACCGGCGCGCGCGCGCGCGCGCGCGCGCG 720  
DB 661 ACCTGTTCCAGTCCGGGACATCAACCGGCGCGCGCGCGCGCGCGCGCGCGCG 720

QY 721 GTGCAATCGCGGACATGCTTGGCTGACACCTTCCGATGATCTTGGCTACTACAC 780  
DB 721 GTGCAATCGCGGACATGCTTGGCTGACACCTTCCGATGATCTTGGCTACTACAC 780  
QY 781 GCGCTGAGCG 840  
DB 781 GCGCTGAGCG 840  
QY 841 AAGAACTGCG 900  
DB 841 AAGAACTGCG 900  
QY 901 GACATCGGATGACGCTTCAAGAGATGACCGGAGCGCGCGCGCGCGCGCGCGCG 960  
DB 901 GACATCGGATGACGCTTCAAGAGATGACCGGAGCGCGCGCGCGCGCGCGCGCG 960  
QY 961 TTCGCTATGCG 1020  
DB 961 TTCGCTATGCG 1020  
QY 1021 GAGCTGCGCGGACATGACGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 1021 GAGCTGCGCGGACATGACGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
QY 1081 AATGTCGATGCTCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
DB 1081 AATGTCGATGCTCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
QY 1141 ATGTCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200  
DB 1141 ATGTCGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200  
QY 1201 ATCATCGCGCAAC 1212  
DB 1201 ATCATCGCGCAAC 1212

RESULT 6  
US-10-251-078-15  
Sequence 15, Application US/10251078  
Publication No. US20030119084A1  
GENERAL INFORMATION:  
APPLICANT: Roche Diagnostics GmbH  
APPLICANT: Shao, Zhixin  
APPLICANT: Schmuck, Rainer  
APPLICANT: Kratzsch, Peter  
APPLICANT: Kenkies, Janet  
APPLICANT: Weisner, Harald  
TITLE OF INVENTION: Variants of an Erwinia-type creatinase  
FILE REFERENCE: 20981 EP  
CURRENT APPLICATION NUMBER: US/10/251,078  
CURRENT FILING DATE: 2002-09-20  
NUMBER OF SEQ. ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ. ID NO. 15  
LENGTH: 1215  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: variant CT2m28  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1215)  
US-10-251-078-15

Query Match 80.6%; Score 976.8; DB 15; Length 1215;  
Best Local Similarity 87.9%; Pred. No. 1.7e-252;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGAAATGCAACAACGCGAGAAAGATTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGAAATGCAACAATGCGAAGAAATATTCGCG 60

QY 61 TTTTGGATGCGGAGATGACCCGCGCAAAAGAGCTTGGGGTGAATGGCGAAGAC 120  
DB 61 TTTTCCGATGCGGAGATGACCGCGCGCAAGTGAATGCGCGCTGATGGCCGAAGAC 120  
QY 121 AATGTGATGCGGCGCTGTTCACCTTTTATCATCTGATCACTACTATTCGCGCTG 180  
DB 121 GAGCTGAGGCTGGCGTGTTCACCTTCATCATTTGATCACTACTACTGATTCCTG 180  
QY 181 TACTGCTATTTGGAGCGCAAGTACGGCATGTGATGACCAACAACCGCCAGCAT 240  
DB 181 TACTGCTATTTGGCGCGCAAAATACGGCATGTGATGACCAAGACCATTCGACCATC 240  
QY 241 TGGGCGGCGATGAGCGGCGCGAGCCCTGCGCGCGGAGTTTGGGGGCAACAATCACTTAC 300  
DB 241 TGGGCGGCGATGAGCGGCGCGAGCCCTGCGCGCGGAGTTTGGGGGCAACAATCACTTAC 300  
QY 301 ACCGACTGCGCGCGCAATTTCTATTCGCGCGTGGCGCACTGACCAACGCGCGCGCAAG 360  
DB 301 ACCGACTGCGCGCGCAATTTCTATTCGCGCGTGGCGCACTGACCAACGCGCGCGCAAG 360  
QY 361 CGCATGCGCATGAGTTTGAACAACGTCATTCGACTTCGCGCGGCGCATGAGAGAGCC 420  
DB 361 CGCATGCGCATGAGTTTGAACAACGTCATTCGACTTCGCGCGGCGCATGAGAGAGCC 420  
QY 421 CTACCGGCGCTGCACTTGTGACATCAGCGCGCTGATGTGATGAGCGACATTCAG 480  
DB 421 CTACCGGCGCTGCACTTGTGACATCAGCGCGCTGATGTGATGAGCGACATTCAG 480  
QY 481 TCGCTCGAAGAGCAGAGCTGATTCGCGAGAGCGCGCGCGTGTGACGTCGCGCGCGCG 540  
DB 481 TCGCTCGAAGAGCAGAGCTGATTCGCGAGAGCGCGCGCGTGTGACGTCGCGCGCGCG 540  
QY 541 GCGTGGCGCGCTGCCATCAAGCGCGCGTGGCGAGCATGAAGTGGCGATGCCACAC 600  
DB 541 GCGTGGCGCGCTGCCATCAAGCGCGCGTGGCGAGCATGAAGTGGCGATGCCACAC 600  
QY 601 AATGGATGATCGCGGAGATGCGCAAAATGTTCCCTTGTGGAGTGAATGAGCACTGG 660  
DB 601 AATGGATGATCGCGGAGATGCGCAAAATGTTCCCTTGTGGAGTGAATGAGCACTGG 660  
QY 661 ACCTGTTTCAGTCCGCGCATCAACCGAGCGCGCGCAATCCGCTCAACAACCGCATC 720  
DB 661 ACCTGTTTCAGTCCGCGCATCAACCGAGCGCGCGCAATCCGCTCAACAACCGCATC 720  
QY 721 GTGCAATCGCGGCACTCTTGTGCTCAACACCTTCCGATATCTTGGCTACTACAC 780  
DB 721 GTGCAATCGCGGCACTCTTGTGCTCAACACCTTCCGATATCTTGGCTACTACAC 780  
QY 781 GGGCTGGAGCGGACCGTGTTCGCGACATGTGAGTGAAGCGGCGTGGCACTTGGAG 840  
DB 781 GGGCTGGAGCGGACCGTGTTCGCGACATGTGAGTGAAGCGGCGTGGCACTTGGAG 840  
QY 841 AAGAAGTGGCGCGCATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900  
DB 841 AAGAAGTGGCGCGCATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900  
QY 901 GACATGCGCATGAGTCAAGATGTAACCGGAGTGGGACCTGTGTAAGTACCGCTCC 960  
DB 901 GATATGCGCGCTGAACTCAACGAGATGTACCGGAGTGGGATCTGTGTAAGTACCGCTCC 960  
QY 961 TTTGGCTATGAGGCACTCTTGGCGGCTGTGCTCACTAATGAGTGGCGCGCGCTG 1020  
DB 961 TTTGGCTATGAGGCACTCTTGGCGGCTGTGCTCACTAATGAGTGGCGCGCGCTG 1020  
QY 1021 GAGTGGCGGAGGACATCGACACCGAGCTGAAGCCGCGCATGTGGTCTCAATGAGCGG 1080  
DB 1021 GAGTGGCGGAGGACATCGACACCGAGCTGAAGCCGCGCATGTGGTCTCAATGAGCGG 1080  
QY 1081 AATGTGATGCTCCGAGGCGCATGCGCGGCTGGCGCGGCTTATGCGAGCAGCATCTG 1140  
DB 1081 AATGTGATGCTCCGAGGCGCATGCGCGGCTGGCGCGGCTTATGCGAGCAGCATCTG 1140  
QY 1141 ATCGTGGGGAAGAGGCTGCGAGAACATCACCGGCTTCCGTTGGTCCGGAACAACACC 1200

DB 1141 ATCGTGGGGAAGAGGCTGCGAGAACATCACCGGATTCCTTCCGCGCTGAGACAAAC 1200  
QY 1201 ATCATCCGCAAC 1212  
DB 1201 ATCATCCGCAAC 1212

RESULT 7  
US-10-251-078-21  
; Sequence 21, Application US/10251078  
; Publication No. US20030119084A1  
; GENERAL INFORMATION:  
; APPLICANT: Roche Diagnostics GmbH  
; APPLICANT: Shao, Zhixin  
; APPLICANT: Schumuck, Rainer  
; APPLICANT: Kratzsch, Peter  
; APPLICANT: Kemlries, Janet  
; APPLICANT: Weisner, Harald  
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase  
; FILE REFERENCE: 20981 EP  
; CURRENT APPLICATION NUMBER: US/10/251,078  
; CURRENT FILING DATE: 2002-09-20  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 1212  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: variant C7ed7  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1212)  
US-10-251-078-21

Query Match 80.5%; Score 975.2; DB 15; Length 1212;  
Best Local Similarity 87.8%; Pred. No. 4.7e-252;  
Matches 1064; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGATGAATGAGCAACGCGCGAAGAAATTTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGATGAATGAGCAATGATGAGAAAGAAATATTCGCC 60  
QY 61 TTTTGGATGCGGAGATGACCCGCGCAAAAGAGCTTGGGGTGAATGGCGAAGAC 120  
DB 61 TTTTCCGATGCGGAGATGACCGCGCGCAAGTGAATGCGCGCTGATGGCCGAAGAC 120  
QY 121 AATGTGATGCGGCGCTGTTCACCTTTTATCATCTGATCACTACTATTCGCGCTG 180  
DB 121 GAGCTGAGGCTGGCGTGTTCACCTTCATCATTTGATCACTACTACTGATTCCTG 180  
QY 181 TACTGCTATTTGGAGCGCAAGTACGGCATGTGATGACCAACAACCGCCAGCAT 240  
DB 181 TACTGCTATTTGGCGCGCAAAATACGGCATGTGATGACCAAGACCATTCGACCATC 240  
QY 241 TGGGCGGCGATGAGCGGCGCGAGCCCTGCGCGCGGAGTTTGGGGGCAACAATCACTTAC 300  
DB 241 TGGGCGGCGATGAGCGGCGCGAGCCCTGCGCGCGGAGTTTGGGGGCAACAATCACTTAC 300  
QY 301 ACCGACTGCGCGCGCAATTTCTATTCGCGCGTGGCGCACTGACCAACGCGCGCGCAAG 360  
DB 301 ACCGACTGCGCGCGCAATTTCTATTCGCGCGTGGCGCACTGACCAACGCGCGCGCAAG 360  
QY 361 CGCATGCGCATGAGTTTGAACAACGTCATTCGACTTCGCGCGGCGCATGAGAGAGCC 420  
DB 361 CGCATGCGCATGAGTTTGAACAACGTCATTCGACTTCGCGCGGCGCATGAGAGAGCC 420  
QY 421 CTACCGGCGCTGCACTTGTGACATCAGCGCGCTGATGTGATGAGCGACATTCAG 480  
DB 421 CTACCGGCGCTGCACTTGTGACATCAGCGCGCTGATGTGATGAGCGACATTCAG 480  
QY 481 TCGCTCGAAGAGCAGAGCTGATTCGCGAGAGCGCGCGTGTGACGTCGCGCGCGCG 540

Db 481 TCGCTCGAAGACAGAAAGCTGATCCGAGGGTCCCGTATCTCGACACTCCGCGTCC 540  
Qy 541 GCCTGCGCGGCTGCCATCAAGGCGGCGTCCGAGCATGAAGTGGCGATCCCAAC 600  
Db 541 GCTCGCTGGCGCGGCTGCAAGGCGGCGTCCGAGCAAGAGTGGCGATCCCAAC 600  
Qy 601 AATGGATGATCCGCGAGATCCGCAATGTTCCCTTGTGTGAGTGTATGGAACCTGG 660  
Db 601 AATGGGATGATCCGCGAGATCCGCAAGTGTTCCTTGTGTGAGTGTATGGAACCTGG 660  
Qy 661 ACCTGTTCCAGTCCGCGCATCAACCGAGCGGCGCAATCCGCTGCAACCGCATC 720  
Db 661 ACCTGTTCCAGTCCGCGCATCAACCGAGCGGCGCAATCCGCTGCAACCGCATC 720  
Qy 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTCAAC 780  
Db 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTCAAC 780  
Qy 781 GCGCTGAGGCGGAGCTGTTCTGCGAATCTGATGAGCGCAAGCTTGCATCTGGAG 840  
Db 781 GCGCTGAGGCGGAGCTGTTCTGCGAATCTGATGAGCGCAAGCTTGCATCTGGAG 840  
Qy 841 AAGAACTGGCGGCGATCCGCGGCGCTGAGCTGATCAACCGGCGCGGCTGCAAG 900  
Db 841 AAGAACTGGCGGCGATCCGCGGCGCTGAGCTGATCAACCGGCGCGGCTGCAAG 900  
Qy 901 GACATCCGCGATCCGCGGCGCTGAGCTGATCAACCGGCGCGGCTGCAAG 960  
Db 901 GACATCCGCGATCCGCGGCGCTGAGCTGATCAACCGGCGCGGCTGCAAG 960  
Qy 961 TTGCGCTATGCGCATCTCTTCCGCGTGTGCTGCTACTACTAGTGGCGAGCGGCTG 1020  
Db 961 TTGCGCTATGCGCATCTCTTCCGCGTGTGCTGCTACTACTAGTGGCGAGCGGCTG 1020  
Qy 1021 GAGCTGCGGAGGAGATCGACACGAGCTGAAGCGCGGATGTTGCTTCCATGAGGCG 1080  
Db 1021 GAGCTGCGGAGGAGATCGACACGAGCTGAAGCGCGGATGTTGCTTCCATGAGGCG 1080  
Qy 1081 ATGTGATGCTGCTCGGAGGAGATCCCGGTGCGGCGGCTATCGGAGACGACATCTG 1140  
Db 1081 ATGTGATGCTGCTCGGAGGAGATCCCGGTGCGGCGGCTATCGGAGACGACATCTG 1140  
Qy 1141 ATGCTGCGGAGAGAGCTGTCGAGAAATCACTACCGCTTCCGTTGCTGCGAGCAAC 1200  
Db 1141 ATGCTGCGGAGAGAGCTGTCGAGAAATCACTACCGCTTCCGTTGCTGCGAGCAAC 1200  
Qy 1201 ATCATCCGCAAC 1212  
Db 1201 ATCATCCGCAAC 1212

## RESULT 8

US-10-251-078-19

; Sequence 19, Application US/10251078

; Publication No. US20030119084A1

; GENERAL INFORMATION:

; APPLICANT: Roche Diagnostics GmbH

; APPLICANT: Shao, Zhixin

; APPLICANT: Schneck, Rainer

; APPLICANT: Kratzsch, Peter

; APPLICANT: Kenkies, Janet

; APPLICANT: Weisner, Harald

; TITLE OF INVENTION: Variants of an Erwinia-type creatinase

; FILE REFERENCE: 20981 EP

; CURRENT APPLICATION NUMBER: US/10/251,078

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 19

; LENGTH: 1212

; TYPE: DNA

; ORGANISM: Artificial Sequence

FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:variant Ctsd2  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1) .. (1212)  
US-10-251-078-19

Query Match 80.3%; Score 973.6; DB 15; Length 1212;  
Best Local Similarity 87.7%; Pred. No. 1.3e-251;  
Matches 1063; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACATGTTGACAGTATGAAATGGGACAAAGCGGAGAAAGATTATTCGCG 60  
Db 1 ATGACTGACGACATGTTGACAGTATGAAATGGGACAAAGTGTGAGAAAGATTATTCGCG 60  
Qy 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGCAAGTTCGCGGCTGATGGCCAAAG 120  
Db 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGCAAGTTCGCGGCTGATGGCCAAAG 120  
Qy 121 AATGTCGATGCGGCGGCTGTTCACTCTTATCACTGATCAACTATATTCGGCTGCTG 180  
Db 121 AATGTCGATGCGGCGGCTGTTCACTCTTATCACTGATCAACTATATTCGGCTGCTG 180  
Qy 181 TACTGCTATTTCCGACGCAAGTACGGAATGATGATGACCAACCAAGCCACGACGATT 240  
Db 181 TACTGCTATTTCCGACGCAAGTACGGAATGATGATGACCAAGCCACGACGATT 240  
Qy 241 TCGGCGGCGCATGACGCGGCGCAGCCCTGCGCGCGCACTTCCGCGCAACATCACTAC 300  
Db 241 TCGGCGGCGCATGACGCGGCGCAGCCCTGCGCGCGCACTTCCGCGCAACATCACTAC 300  
Qy 301 ACCGACTGGCGCGCGGACATTTCTATTCGCGCGCTGCGCGCTGACCAAGCGGCGCAAG 360  
Db 301 ACCGACTGGCGCGCGGACATTTCTATTCGCGCGCTGCGCGCTGACCAAGCGGCGCAAG 360  
Qy 361 CGCATCCGCGCATGAGTTGACACGATCAATTCGACCTTCCGCGCGCACTGAGGAGCG 420  
Db 361 CGCATCCGCGCATGAGTTGACACGATCAATTCGACCTTCCGCGCGCACTGAGGAGCG 420  
Qy 421 CTACCGGCGGCTGACCTTCTGATGACATGACGAGCCCTGATGATGAGTGGACCATCAAG 480  
Db 421 CTACCGGCGGCTGACCTTCTGATGACATGACGAGCCCTGATGATGAGTGGACCATCAAG 480  
Qy 481 TCGCTCGAAGAGCAAGCTGATCCGGAAGCGCGCGCTGTGTGATGAGTGGCGCGCG 540  
Db 481 TCGCTCGAAGAGCAAGCTGATCCGGAAGCGCGCGCTGTGTGATGAGTGGCGCGCG 540  
Qy 541 GCGTGGCGGCTGACATCAAGGCGGCGGTCGCGAGCATGAAGTGGCATGCGCGCAAC 600  
Db 541 GCGTGGCGGCTGACATCAAGGCGGCGGTCGCGAGCATGAAGTGGCATGCGCGCAAC 600  
Qy 601 AATGCAATGATCCGCGAGATGCGCAAAATCGTTCCCTTCTGTGAGCTGATGAGCACTGG 660  
Db 601 AATGCAATGATCCGCGAGATGCGCAAAATCGTTCCCTTCTGTGAGCTGATGAGCACTGG 660  
Qy 661 ACCTGTTCCAGTGGGCGATCAACCGAGCGGCGCCCAAAATCCGCTGATCAACCGCATC 720  
Db 661 ACCTGTTCCAGTGGGCGATCAACCGAGCGGCGCCCAAAATCCGCTGATCAACCGCATC 720  
Qy 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACTTCCGATGATCTTGGGCTACTACAC 780  
Db 721 GTGCAATCCGCGGAGATCTTTCGCTCAACACTTCCGATGATCTTGGGCTACTACAC 780  
Qy 781 GCGCTGAGGCGACAGCTGTTCTGCGAACATGATGATGAGCGGCGGCTGATGAGGAG 840  
Db 781 GCGCTGAGGCGACAGCTGTTCTGCGAACATGATGATGAGCGGCGGCTGATGAGGAG 840  
Qy 841 AAGAACTGGCGGCTGCAATCCGCGGCGCTGAGCTGATCAAGCGGCGGCGCGCTGCAAG 900  
Db 841 AAGAACTGGCGGCTGCAATCCGCGGCGCTGAGCTGATCAAGCGGCGGCGCGCTGCAAG 900  
Qy 901 GACATCCGCGATCCGCGGCGCTGAGCTGATCAACCGGCGGCGGCTGCAAG 960  
Db 901 GACATCCGCGATCCGCGGCGCTGAGCTGATCAACCGGCGGCGGCTGCAAG 960





QY 265 CCTGAGCGCCGAGCTTGGGCGCAACATCACTACCGGATGGCGCCGCAAAATTTC 324  
DB 1047 AGCATGCGCTTCTACTTGGCCGACGAGACTTCTTCAAGCGCGCTCGCGGATTC 988  
QY 325 TATCGCGCGCTGCGCGAGCTGACACACGCGCGCCAAAGCGCATCGGATCGAGTTCAGCAC 384  
DB 987 GGGCTCGCCCTGAGTGGGCAACGGAAGGTGGCGAGCTCGCCCGGACCCGCGCAC 928  
QY 385 GTCAATCTGCACTTCGCGCGCGAGCTGAGAAAGCTTACCGGCGCTGCACTTCTGCAAC 444  
DB 927 GCGCGGCTCTGGGCTCGGACAAAGTGGAGGCTTACTACCGGAGATGAGTCAAGTAC 868  
QY 445 ATGACCAAGCTTCGATGTGATGTGGACATCAAGTGGCTGGAAGAGAGAAAGTGTATC 504  
DB 867 GTGCGTGTGGGAGCAAGAGTCTCTCCCGCGCGCGCGCTGTCTGACAGGCAATGAGAAC 808  
QY 505 CGCGAAGGCGCGCGCGCTGTGACGTGCGCGCGCGCGCTGTGCGCGCTGCAATCAAGGCC 564  
DB 807 GTGACAGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTGCAAGTGTCAAGCGCGCTC 748  
QY 565 GCGCGTGGCGGAGCATGAAAGTGGCGATCGGCGCACCAATGCGATGATCGCGAGATCGCC 624  
DB 747 AAGATGAGACGCGGTGACATCTCTGCGCGCGCTGCGCGCGCTGTTGAGGAGACCGCGCC 688  
QY 625 AAATGCTTCCCTTCTGTGAGCTGATGACATCTGGAAGTGTCTTCAAGTGGGCAATCAAC 684  
DB 687 GCGAGTGTGGCGCGCGCGAGTTCGTGCGCGCGCGCGCGCGCTGTCTGCGCAAGCTTC 628  
QY 685 ACCGAGCGCGCGCGCATCCGGTCAACCAACCGCATGTGCAATCGCGCGCAATCTCTTTCG 744  
DB 627 TACCCCTACTTGGCGCTGAGTACAGGACGCGCGCATGACTTGAATCAAGCGCTTTC 568  
QY 745 CTCAACACCTTCCCGATGATCTTGGCTTCTACACCGCGCTGAGAGCGAGCTGTTTCG 804  
DB 567 CAGCCGAGCGCGCGCGCGTCAAGGACCGCGCGCGCGCGCGCTGTGATGACGAACTCTTTC 508  
QY 805 GACCATGTGATGAGCGCGCGCTGCAATCTGAGAGAAAGTGGCGCTGATGCGCGC 864  
DB 507 GACGCGATGTGAGCGCGCGCTGCGCGCGCGCTGAGAACAGGCGCGCGCGCGCGCTTC 448  
QY 865 GGGCTCGAGCTGATCAAGCGCGCGCGCGCGCTGCAAGGAC 903  
DB 447 GACGTGTGTGTGCGAGACGCGGATGGCGCTGCGCGGAC 409

## RESULT 11

US-10-437-963-38426  
; Sequence 38426, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; PRIORITY FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 38426  
; LENGTH: 1483  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_42062C.1  
US-10-437-963-38426

Query Match 5.8%; Score 69.8; DB 17; Length 1483;

Best Local Similarity 44.9%; Pred. No. 1,1e-08;  
Matches 266; Conservative 0; Mismatches 327; Indels 0; Gaps 0;

QY 587 CGATCGCCACCAACCAATGCGATGATTCGCGAGATCGCAAAATGTTCCCTTGTGGAGC 646  
DB 309 CGAGACCGCGCGAGCGGAGATTCGCGCGCTTCACTCGCGGAGTACGACCTTC 368  
QY 647 TGATGACACTTGAACCTGTTTCCAGTGGGATCAACACGAGCGCGCGCAATTCGG 706  
DB 369 TCCGAGACTTACCCCGAGGCTACTTCAACGAGCGCGCTTGGAGAGGCGGAG 428  
QY 707 TCACCAACCGATGCGTGAATTCGCGGCAATCTTTCGCTCAACACTTCCGATGATCT 766  
DB 429 ATGACACGAGGATGCGCGGCAACGATGCGCGCTTTCGACCGCTTGGAGTACT 488  
QY 767 TCGGCTACTACCGCGCTGAGCGCAAGCTTCTGCGACCATGTGATGAGCGCGC 826  
DB 489 GCGCGGCTTACCGCGCGGCTGTTGCGCGCGCGCGCGCTTGTGACGCGCGCTCG 548  
QY 827 TCGACATTTGGAGAGAAAGTGGCGCTGCTGCTGCGCGCGCGCTGAGCTGATCAAGCGG 886  
DB 549 ACATCGCCATCACTGTGCGCGCGCGCATGACACGCGCTGCGCTGCAAGGCAAGCGCT 608  
QY 887 GCGCGCGCTGCAAGGACATCGCGCATTCGAGCTCAACGAGATGTACCGCGAGTGGAGCTGC 946  
DB 609 TCTGCTACGTAACGACATCTGCTGCTGCGCATCAAGAGCTCTTGGGACCTTCAGCGAG 668  
QY 947 TGAAGTACCGCTCTTCTGCGCTATGAGCGCATCTCTTCTGCGCTGTGCTGCTGCTGCTC 1006  
DB 669 TCACTTACGTGACATGACCGCGCACACGCGCGAGGCTGACAGACCTTCTTGGACT 728  
QY 1007 GCGAGCGCGGCTGAGTGGCGGAGGACATGACACCGAGCTGAAGCGCGGAGTGG 1066  
DB 729 CCAACCGGCTATGACGCTGCTGCTTCCACCGGTAGGAGAAATCAAGCTTCCACAGGACT 788  
QY 1067 TCTCAGTACCGGATGATGCTGCGCGGAGGAGCATCCGCTGCGCGGCTTATCGCG 1126  
DB 789 TCTTCCGCGCTCGCGCGCATCAACGAAATCGGAGAGCGCGCGGACGACATCAAGG 848  
QY 1127 AGCAGCATCTGATCTGCGGAGAGAGCGGTGCGGAGACATCAAGCGCTTC 1179  
DB 849 TGAAGTCCCTTGGAGCGCGCGCTGCGCGAGAGCTTACAGAGCTGCTTC 901

## RESULT 12

US-10-156-761-1  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HAYTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIORITY FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIORITY FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4187715)  
; OTHER INFORMATION: a, t, c, g, other or unknown



Oy	842	AGAAACGAGGCGGCGTCAATCGCGCGGGGCTCGAGCTGATCAAGCCGGGCGCGCGCTGCAAG	901
Db	18966	TCGAATCTGGGCGCCGAGCGGCTGCTGTGTCGCGCATAGCGCCGAGAAATCTTGGCCGAGAGGG	190255
Oy	902	ACATTCGCATCGAGCTCAACGAGATATGACGCGAGTGGAGCTGATGAAGTACCGCTCT	961
Db	19026	CCGCGCGCGGTGCTCCGCTGTGCGGCAAGAACCGCCCCGAGGAGCTCTCCGCGCTACCTGGCC	190855
Oy	962	TCGGCTATGCGSCACTCTTCGGCGCTGTGTGACATTAACGATCGCGAGCGCGCGGTGG	1021
Db	19086	TGGCGCGCGGCCCAAGTCCGCGGGGCTCAACGGATCCGCTGTGGGCGGCGCTTTCGAGCGGACCG	19145
Oy	1022	AGCTGCGCGAGGACATGACACGAGCTGAAGCCCGGCGATGTGTGCTTCATGAGACCGA	1081
Db	19146	GCGCGCGCGCGCGCGCACTTGCCCACTAACCTTCCTTCAGCACGACGGATCTGGCCGACCG	19205
Oy	1082	TGTGTGATGTCTCCGAGGAGGCAATGCCCGGTGTCCGCGGCTATCCCGAGCACGACATCTTGA	1141
Db	19306	CGGCGCGCGCGCGCCAGAGAGGTCAACCCGCGCGGAGACTGGCGCGCGCGGACCAACCGCGTGC	19265
Oy	1142	TGCTTCGGGGAAGGAGCGTTCGCGAAGAACATCAACCGGCTTCGCCGTTTC	1185
Db	19266	TCGGCGCGCAACGTGTGAATCTGCGCGACGCGGGCGCGCTACTTGTGTTTC	19309

### RESULT 14

```

Sequence 35 Application US/10203295
Publication No. US20040115762A1
GENERAL INFORMATION:
APPLICANT: Zolchev, Sergey Borisovich
APPLICANT: Sekurova, Olga Nikolayivna
APPLICANT: Fjaeravik, Espen
APPLICANT: Brautaset, Trygve
APPLICANT: Strom, Arne Reidar
APPLICANT: Valla, Svein
APPLICANT: Ellingse, Trond Erling
APPLICANT: Sletta, Havard
APPLICANT: Gulliksen, Ole-Martin
TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
FILE REFERENCE: 1181-265
CURRENT APPLICATION NUMBER: US/10/203,295
CURRENT FILING DATE: 2003-05-19
PRIOR APPLICATION NUMBER: PCT/GB 01/00509
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: GB 0002840.7
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: GB 0008786.6
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: GB 0009387.2
PRIOR FILING DATE: 2000-04-14
SOFTWARE: PatentIn version 3.0
SEO ID NO 35
LENGTH: 125401
TYPE: DNA
ORGANISM: Streptomyces noursei ATCC 11455
US-10-203-295-35

```

	Query Match	Similarity	5.7%	Score 68.8	DB 17	Length 125401
	Best Local	Similarity 42.8%	Pred. No. 4.5e-08			
	Matches 404	Conservative 0	Mismatches 517	Indels 3	Gaps 1	
Oy	245	CCGCGATCGAGCGGGCCAGCCCTGGGCGCGAGACTTCGGCGACACATCACTTACACCG	304			
Db	78627	CCGCGCTGTTCGCGCTGAGGTGCGCCCTTACCGCGCTGTGCGGTCCCTGGCGTACCC	78686			
Oy	305	ACTGCGCGCGGACAAATTTCTATCGCGCGGCGCCAGCTGACCAACGCGCCACAGCGCA	364			
Db	78687	CCGACTTCGTGCGGGCGCATCTCATCGGCAGCTCGCGCGCGGACGTGCGGSGTCC	78746			
Oy	365	TCGGCATCGATTGACCAAGTCATCTGACTTCGCGCGACGCTCGAGAGAGCCCTTAC	424			

Db	78747	TTCTCCTTCGAAGACGGCTTGACCTCTGTGCGCCGCGCCGCTTATGAGGCCCTTGC	78806
Qy	425	CGGCGCTGCATCTTCGTGACATCAGCCAGCCCTCGATGTGGATNCGACCATCAATGTCG	484
Db	78807	CGGCGCGCGCGCGGATGCTGTGGGATTCGGGCGCACCGAGGACGAGGTACACCCCCACCTCA	78866
Qy	485	TCGAGAGCAGAAAGCTGATTCGCGAAGCGGCCGCGTGTGATGCTGGCGCGCGGCT	544
Db	78867	CGAGACACGCTTCGATCGCGCGCGGTACCGGGGCCACTTCGTCGTGTGTCGCGCACCG	78926
Qy	545	GGCGCGCTCCATCAAGGCGCGGCTCCGAGATGAATGTGGGATTCGCCACCAATG	604
Db	78927	AGGAAGCCGTTCGCGCGGATGGGGGCGCGCTTCAACGCCACAGAACCGAAGCCACCGGC	78986
Qy	605	CGATGATTCGCGGAGATTCGCCAAATCGTTCCTTCGTGAGCTGATGAGACACTGACCT	664
Db	78987	TGCGGGTACAGCCACGCTTTCACCTCGCGCTCATGAGCCGGAATCTGGCGAATTCGCG	79046
Qy	665	GGTTCAGTTCGGGCGATCAACACCGAGCGCGGCAATTCGG--TCACCAACCGCATCG	721
Db	79047	CGGTGCGCGCGGCGCTGACCTACACGAGCGCGGACATCCGGTCTCTCCAACTTACCG	79106
Qy	722	TGCAATCGGCGCAATCTTTGCTCAACACTTCCGATGATCTTGGGCTACTACACCG	781
Db	79107	GCACCGTTCGCGCGCGGTGCGGACCTGTGTCTCCGCACTATGGGTCGCGACGTCGCG	79166
Qy	782	CGCTGGAGGCAACGCTTCTGTGCAACATGTGATGACGCGCAGCCTGCACATCTGGAGA	841
Db	79167	AGGCGGTCCGCTTTCGCGCCAGCGGCTACCGGCTTCAACGACCGCGGCTGACACAGCTCG	79226
Qy	842	AGAACGTGGCGGTGCATTCGCGCGCGGCTCGAGCTGATCAAGCCGGGCGCGCTGCAAG	901
Db	79227	TGGAATCGGCGCGGACGCGGCTGTGTGCGCCATGGCGCCAGAACTCTCGCGGACGCG	79286
Qy	902	ACATGGCCATCGAGCTCAACGAGATATATCCGCGAGTGGACCTGTCTGAATGACCGCTCT	961
Db	79287	CGCGCGCTGTGCGGCTGTGCGCACAGAACCGCCCGAGGAGCTTTCGCGCTACCGGC	79346
Qy	962	TGCGGATATGCGCACTCTTCGCGGCTGTGCGCACTACTACGATCGCGGAGGCGCGGTG	1021
Db	79347	TGGCCCGGCCCACTGCTCGCGGCGGTACGATCCGCTGTGGCGCGGCTTTTGACGCGCACG	79406
Qy	1022	AGCTGCGCGAGGACATTCGACACCGAGCTGAAGCCCGGACGATGTGTCTTCATGAGCGGA	1081
Db	79407	GGGCGGCGCGCGCGGACCTGACCACCTAACCCCTTCAGACACAGCGGTTTCGGCGACCG	79466
Qy	1082	TGGTATGTCTCCGAGGGGCAATGCCGTGCGCGGCGGCTATTCGAGACAGACATCTTGA	1141
Db	79467	CGGCGCGCGCGCCACGAGGACGTCAACGCGCGCGGGACTGGGCGCGCGCGACACACCGCTGC	79526
Qy	1142	TGCTGGGGAGGACGATGCCGAGAAACATCACCGGCTTCCCGTTCC	1185
Db	79527	TGGGCGCACCGTGCACCTGCGGACGAGGCGCGGCTACTTGTTC	79570

## RESULT 15

```

US-10-437-963-57011/C
; Sequence 57011, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 14:49:20 ; Search time 5273 Seconds

(without alignments)  
10869.546 Million cell updates/sec

Title: US-10-807-228-2

Perfect score: 1212  
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACCATCATCGCGAAC 1212

## Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hhg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	6	AR100122 Sequence
2	1210.4	99.9	1212	6	E12280 DNA encodin
3	1210.4	99.9	1212	6	E13584 Alkali genes
4	1210.4	99.9	1212	6	E16405 Alkali genes
5	1210.4	99.9	1212	6	BD017699 Gene enco
6	1210.4	99.9	1212	6	BD105662 Stable cr
7	1205.6	99.5	1215	1	AB016788 Alkali gen
8	1184.8	97.8	1215	6	E11155 gDNA encodi
9	981.6	81.0	1215	6	AX721933 Sequence
10	978.4	80.7	1215	6	AX721949 Sequence
11	978.4	80.7	1215	6	AX721941 Sequence
12	976.8	80.6	1212	6	AX721943 Sequence
13	976.8	80.6	1212	6	AX721945 Sequence
14	976.8	80.6	1215	6	AX721947 Sequence
15	975.2	80.5	1212	6	AX721953 Sequence
16	973.6	80.3	1209	6	AX721951 Sequence
17	640.2	52.8	1202	6	E10020 DNA encodin
18	640.2	52.8	1723	1	FVBCR1 Flavobacter
19	640.2	52.8	1956	1	FVBCRE Flavobacter

20	635.6	52.4	1212	6	A10619 Recombinant
21	635.6	52.4	1212	6	E00904 gDNA encodi
22	635.6	52.4	1212	6	I01844 Sequence 2
23	634	52.3	1212	6	A01506 Recombinant
24	579.4	47.8	300029	1	AE016787 Pseudomon
25	568.8	46.9	3774	1	AF170566 Pseudomon
26	567.2	46.8	1810	1	AF072304 Pseudomon
27	564.6	46.6	1134	6	E01576 DNA sequenc
28	457.8	37.8	1233	6	E01828 Genomic DNA
29	457.8	37.8	1584	1	BACCR2
30	374.4	30.9	8127	1	AB007122 Arthrobac
31	372.8	30.8	1236	6	E17219 gDNA encodi
32	84.4	7.0	303855	1	AE017230 Mycobacte
33	78.2	6.5	291000	1	SC039105 Streptomy
34	76	6.3	321250	1	SC039111 Streptomy
35	75.8	6.3	302998	1	AE016921 Chromobac
36	75.4	6.2	2783	8	AK119861 Oryza sat
37	75.4	6.2	4740	8	AF271357 Oryza sat
38	75.4	6.2	146841	8	AP003629 Oryza sat
39	73.2	6.0	125020	1	AF429315 Homo sapi
40	73.2	6.0	295500	1	SC039128 Streptomy
41	73	6.0	295150	1	SC039126 Streptomy
42	72.6	6.0	110000	1	RME591985_04 Continuation (5 of
43	72.6	6.0	346510	1	AP003011 Mesorhizo
44	72.4	6.0	786	6	AX928131 Sequence
45	72.4	6.0	786	6	AX928133 Sequence

## ALIGNMENTS

RESULT 1  
AR100122  
LOCUS AR100122 Sequence 2 from patent US 6080553.  
DEFINITION AR100122  
ACCESSION AR100122  
VERSION AR100122.1 GI:12810570  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1212)  
AUTHORS Sogabe,A., Hattori,T., Nishiy,Y. and Kawamura,Y.  
TITLE Creatine amidinohydrolase, production thereof and use thereof  
JOURNAL Patent: US 6080553-A 2 27-JUN-2000;  
FEATURES  
source  
Location/Qualifiers  
1..1212  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;  
Best Local Similarity 99.9%; Pred. No. 2.6e-145;  
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGACTGACGACATGTTGCACTGATGAATGCGCAACGCGGAGAAAGATTATTCGCGG	60
DB	1	ATGACTGACGACATGTTGCACTGATGAATGCGCAACGCGGAGAAAGATTATTCGCGG	60
QY	61	TTTTCGGATGCGGATGATGACCGCGGCAAAAGACGTTGCGGCTGGATGCGCAAGAC	120
DB	61	TTTTCGGATGCGGATGATGACCGCGGCAAAAGACGTTGCGGCTGGATGCGCAAGAC	120
QY	121	AATGTCATGCGGCGCTGTTCACTTATCACTGATCACTAATTCGCGTGCTG	180
DB	121	AATGTCATGCGGCGCTGTTCACTTATCACTGATCACTAATTCGCGTGCTG	180
QY	181	TACTGCTATTTCGACCGCAAGTACGATGATGATGATGATGATGATGATGATGATGATG	240
DB	181	TACTGCTATTTCGACCGCAAGTACGATGATGATGATGATGATGATGATGATGATGATG	240
QY	241	TCGACCGCATTCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300
DB	241	TCGACCGCATTCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300

```

Db      241 TCGCCGCGCATCGACGGCGGCGCAGCCCTGGCGCGGAGCTTTCGGCGCAACAATCACTAC 300
Qy      301 ACCGATTCGGCGCGCGCAATTTCTATTCGGCGCGCTGGCGCACTGACACAGCGGCGCAAG 360
Db      301 ACCGATTCGGCGCGCGCAATTTCTATTCGGCGCGCTGGCGCACTGACACAGCGGCGCAAG 360
Qy      361 CGCATCGGCGCATCGAGTTTCGACGACGTCATCTCGACTTCGGCGCGGCGGCGAGGCG 420
Db      361 CGCATCGGCGCATCGAGTTTCGACGACGTCATCTCGACTTCGGCGCGGCGGCGAGGCG 420
Qy      421 CTACCGGCGGCTGCACTTCCTGCAATCAGCGACGCTCGATGTGATGCGCAACCATCAAG 480
Db      421 CTACCGGCGGCTGCACTTCCTGCAATCAGCGACGCTCGATGTGATGCGCAACCATCAAG 480
Qy      481 TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      481 TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy      541 GCGTCGCGGCGCTGCGCATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
Db      541 GCGTCGCGGCGCTGCGCATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
Qy      601 AATGCGATGATCCGCGGAGATCCGCAAAATGCTTCCTTCCTGCGAGCTGATGGAACACTTG 660
Db      601 AATGCGATGATCCGCGGAGATCCGCAAAATGCTTCCTTCCTGCGAGCTGATGGAACACTTG 660
Qy      661 ACCTGGTTCAGTCCGCGGCGCATCAACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db      661 ACCTGGTTCAGTCCGCGGCGCATCAACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Qy      721 GTGCAATCCGCGGCGACATCTTCGCTCAACACCTTCGCGATGATCTTCGCTACTACAC 780
Db      721 GTGCAATCCGCGGCGACATCTTCGCTCAACACCTTCGCGATGATCTTCGCTACTACAC 780
Qy      781 GCGCTGAGAGCGGACCGCTGTTCTGCGACCATGTGATGAGCGGACGCTGCAATCTGGAG 840
Db      781 GCGCTGAGAGCGGACCGCTGTTCTGCGACCATGTGATGAGCGGACGCTGCAATCTGGAG 840
Qy      841 AAGAACGTGGCGCGTGCATCGCGCGGCGCTGCGAGCTGATCAACCGGCGGCGGCGGCGG 900
Db      841 AAGAACGTGGCGCGTGCATCGCGCGGCGCTGCGAGCTGATCAACCGGCGGCGGCGGCGG 900
Qy      901 GACATCGCGCATCGAGCTCAACGAGATGTACCGCGAGTGGGACCTGCTGAGATACGCTCC 960
Db      901 GACATCGCGCATCGAGCTCAACGAGATGTACCGCGAGTGGGACCTGCTGAGATACGCTCC 960
Qy      961 TTGGGCTATGAGCACTCTTCGCGGCTGCTGCTGCACTACGATCGGAGCGGCGGCTG 1020
Db      961 TTGGGCTATGAGCACTCTTCGCGGCTGCTGCTGCACTACGATCGGAGCGGCGGCTG 1020
Qy      1021 GAGCTGCGCGGAGGAGATCGACACGAGCTGAGAGCGGCGGAGTGTGTCTCCATGAGAGCG 1080
Db      1021 GAGCTGCGCGGAGGAGATCGACACGAGCTGAGAGCGGCGGAGTGTGTCTCCATGAGAGCG 1080
Qy      1081 ATGTGATGCTGCGCGGAGGAGTGTGCGCGGCTGCGGCGGCTGCGGAGGAGCAATCTG 1140
Db      1081 ATGTGATGCTGCGCGGAGGAGTGTGCGCGGCTGCGGCGGCTGCGGAGGAGCAATCTG 1140
Qy      1141 ATGCTGCGGAGGAGAGCGTGTGCGGAGAAATCAACCGGCTTCCTGCTGCGGAGCAAC 1200
Db      1141 ATGCTGCGGAGGAGAGCGTGTGCGGAGAAATCAACCGGCTTCCTGCTGCGGAGCAAC 1200
Qy      1201 ATCATCGGCAAC 1212
Db      1201 ATCATCGGCAAC 1212

```

```

RESULT 2
E12280
LOCUS      E12280          1212 bp    DNA          linear    PAT 27-APR-1998
DEFINITION DNA encoding Alcaligenes thermostable creatin amidinohydrolase.
ACCESSION E12280
VERSION    E12280.1 GI:3251114

```

```

KEYWORDS   JP 1996308579-A/1.
SOURCE     Alcaligenes faecalis
ORGANISM   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Alcaligenes.
REFERENCE  1 (bases 1 to 1212)
AUTHORS    Sogabe, A., Yamamoto, K. and Kawamura, Y.
TITLE      GENE ENCODING CREATINE AMIDINOHYDROLASE
JOURNAL    Patent: JP 1996308579-A 1 26-NOV-1996;
            TOYOBO CO LTD
COMMENT     OS   Alcaligenes faecalis
            PN   JP 1996308579-A/1
            PD   26-NOV-1996
            PF   16-MAY-1995 JP 1995117283
            PI   SOGABE ATSUSHI, YAMAMOTO KAZUMI, KAWAMURA YOSHINISU PC
            C1   C12N1/5/09, C12N1/21, C12N9/78, (C12N15/09, C12R1:05), (C12N1/21, PC
            C1   C12R1:425),
            PC   (C12N9/78, C12R1:425);
            CC   strandedness: Double;
            CC   topology: linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key
            FT   source
            FT   /clone='TB3581', /organism='Alcaligenes faecalis' FT
            FT   mat_peptide 1..1212
            FT   /product='Alcaligenes thermostable creatin FT
FEATURES
source
1..1212
location/Qualifiers
/organism="Alcaligenes faecalis"
/mol_type="genomic DNA"
/db_xref="taxon:511"
ORIGIN
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 2.6e-145;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGACTGACGACATGTTGACGATGATGAAATGGGACAAAGCGGAGAAAGATTATTCGCGG 60
1 ATGACTGACGACATGTTGACGATGATGAAATGGGACAAAGCGGAGAAAGATTATTCGCGG 60
1 ATGACTGACGACATGTTGACGATGATGAAATGGGACAAAGCGGAGAAAGATTATTCGCGG 60
61 TTTTCGATGCGGAGATGACCGCGCGCCAAACGACGTTGCGGCTGATGAGCGCAAGAAC 120
61 TTTTCGATGCGGAGATGACCGCGCGCCAAACGACGTTGCGGCTGATGAGCGCAAGAAC 120
121 AATGTCGATGCGGCGGCTGTTCACTCTTATACATGCACTCACTACTATTCGCGGCTGCTG 180
121 AATGTCGATGCGGCGGCTGTTCACTCTTATACATGCACTCACTACTATTCGCGGCTGCTG 180
121 AATGTCGATGCGGCGGCTGTTCACTCTTATACATGCACTCACTACTATTCGCGGCTGCTG 180
181 TACTGCTATTTGCGACCAAGTACGGGATGATGATGATGATGATGATGATGATGATGATGAT 240
181 TACTGCTATTTGCGACCAAGTACGGGATGATGATGATGATGATGATGATGATGATGATGAT 240
241 TCGGCGGCGCATCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
241 TCGGCGGCGCATCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
301 ACCGATCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
301 ACCGATCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
361 CGCATCGGCGCATCGAGTTTCGACGACGTCATCTCGACTTCGGCGCGGCGGCGGCGGCGG 420
361 CGCATCGGCGCATCGAGTTTCGACGACGTCATCTCGACTTCGGCGCGGCGGCGGCGGCGG 420
421 CTACCGGCGGCGCTGCACTTCCTGCAATCAGCGACGCTCGATGTGATGCGCAACCATCAAG 480
421 CTACCGGCGGCGCTGCACTTCCTGCAATCAGCGACGCTCGATGTGATGCGCAACCATCAAG 480

```



QY 481 TCGCTGGAAGACAGAGCTGATCCGGAGAGCGCCCGGTGTGTGACGTGGCGCGCG 540  
 DB 481 TCGCTGGAAGACAGAGCTGATCCGGAGAGCGCCCGGTGTGTGACGTGGCGCGCG 540  
 QY 541 GCCTGCGCGGTGCCATCAAGCGCCGTCGCCGAGCATGAAGTGGCGATGCCACACC 600  
 DB 541 GCCTGCGCGGTGCCATCAAGCGCCGTCGCCGAGCATGAAGTGGCGATGCCACACC 600  
 QY 601 AATGCGATGATCCCGGAGATGCGCAATCTCTTCCCTTGTGGAGCTGATGACCTGG 660  
 DB 601 AATGCGATGATCCCGGAGATGCGCAATCTCTTCCCTTGTGGAGCTGATGACCTGG 660  
 QY 661 ACCGTGTTCCAGTGGGGATCAACAACGAGCGCGGCAATCCGGTACCAACCGCATC 720  
 DB 661 ACCGTGTTCCAGTGGGGATCAACAACGAGCGCGGCAATCCGGTACCAACCGCATC 720  
 QY 721 GTGCAATCCGGCGCATCTCTTCCGCTCAACACTTCCCGAGATGTTGGCTACTACAC 780  
 DB 721 GTGCAATCCGGCGCATCTCTTCCGCTCAACACTTCCCGAGATGTTGGCTACTACAC 780  
 QY 781 GCGCTGAGCGGAGCGCTGTTCTGCGACATGTGATGACGCCAGCTTGCATCTGGAG 840  
 DB 781 GCGCTGAGCGGAGCGCTGTTCTGCGACATGTGATGACGCCAGCTTGCATCTGGAG 840  
 QY 841 AAGAACGTGGCGGTGATCGCGCGGGCTGAGCTGATCAAGCGCGCGCGCTGGAG 900  
 DB 841 AAGAACGTGGCGGTGATCGCGCGGGCTGAGCTGATCAAGCGCGCGCGCTGGAG 900  
 QY 901 GACATCCGCATGAGCTCAAGAGATGTACCGCGAGTGGAGACTTGTAAATACCGCTCC 960  
 DB 901 GACATCCGCATGAGCTCAAGAGATGTACCGCGAGTGGAGACTTGTAAATACCGCTCC 960  
 QY 961 TTGCGCTATGAGCCACTCTTCCGCGGTGCTGTGCTGCTACTGATCGGTGCGAGCGCGCTG 1020  
 DB 961 TTGCGCTATGAGCCACTCTTCCGCGGTGCTGTGCTGCTACTGATCGGTGCGAGCGCGCTG 1020  
 QY 1021 GAGCTGCGCGAGGACATGACAGCGAGTGAAGCCCGGATGATGATGATGATGAGCG 1080  
 DB 1021 GAGCTGCGCGAGGACATGACAGCGAGTGAAGCCCGGATGATGATGATGATGAGCG 1080  
 QY 1081 ATGCTGATGCTGCGCGAGGAGATGCCGCTGCGCGGCTATCCGAGACAGATCTCTG 1140  
 DB 1081 ATGCTGATGCTGCGCGAGGAGATGCCGCTGCGCGGCTATCCGAGACAGATCTCTG 1140  
 QY 1141 ATCTGCGGAGAGACGCTGCGGAGAACTACACCGGCTTCCGCTTCCGAGACACACC 1200  
 DB 1141 ATCTGCGGAGAGACGCTGCGGAGAACTACACCGGCTTCCGCTTCCGAGACACACC 1200  
 QY 1201 ATCATCCGCAC 1212  
 DB 1201 ATCATCCGCAC 1212

## RESULT 3

E13584 1212 bp DNA linear PAT 27-APR-1998  
 LOCUS E13584

DEFINITION Alcaligenes faecalis gene for creatine amidohydrolase.  
 ACCESSION E13584

VERSION E13584.1 GI:3252389  
 KEYWORDS JP 1997215494-A/1

SOURCE Alcaligenes faecalis  
 ORGANISM Alcaligenes faecalis

Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 Alcaligenaceae; Alcaligenes.

1 (bases 1 to 1212)  
 Sogabe A., Hattori T., Nishiyama Y., and Kawamura Y.

TITLE NEW CREATINE AMIDOHYDROLASE, ITS PRODUCTION AND ITS USE  
 JOURNAL Patent: JP 1997215494-A 1 19-AUG-1997;

## COMMENT

OS Alcaligenes faecalis  
 PN JP 1997215494-A/1  
 PD 19-AUG-1997  
 PF 13-FEB-1996 JP 1996025435

PI SOGABE ATSUSHI, HATTORI TAKASHI, NISHIYA YOSHIKI, PI  
 KAWAMURA YOSHIHISA  
 PC C12N9/78, C12N15/09, (C12N9/78, C12R1:05), (C12N15/09, C12R1:19);  
 CC strandedness: Double;  
 CC topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..1212  
 FT /strain='E13581'  
 /Location/Qualifiers  
 /organism='Alcaligenes faecalis' PT  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:511'

ORIGIN  
 Query Match 99.9%; Score 1210.4; DB 6; Length 1212;  
 Best Local Similarity 99.9%; Pred. No. 2,66-145;  
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGTGATGAATGACAAACGCGGAGAAAGATTATTCGCG 60  
 DB 1 ATGACTGACGACATGTTGACGTGATGAATGACAAACGCGGAGAAAGATTATTCGCG 60  
 QY 61 TTTTCGATGCGCGGATGACCCCGCGGCAAAAGCACTTCCGCGCTGATGCCAAGAC 120  
 DB 61 TTTTCGATGCGCGGATGACCCCGCGGCAAAAGCACTTCCGCGCTGATGCCAAGAC 120  
 QY 121 AATGTCGATGCGCGGCTGTTCACTCTTATCACTGATCACTAATTCCTGCTGCTG 180  
 DB 121 AATGTCGATGCGCGGCTGTTCACTCTTATCACTGATCACTAATTCCTGCTGCTG 180  
 QY 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACCAACAGCGCAGACGATT 240  
 DB 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACCAACAGCGCAGACGATT 240  
 QY 241 TCGGCGCGGATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 DB 241 TCGGCGCGGATGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 QY 301 ACCGACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
 DB 301 ACCGACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
 QY 361 CGCATCGGCAATGAGTTGACACAGGTCATCTGCACTTCCGCGCGGCGGCGGCGGCGG 420  
 DB 361 CGCATCGGCAATGAGTTGACACAGGTCATCTGCACTTCCGCGCGGCGGCGGCGGCGG 420  
 QY 421 CTACCGGCGGCTGCACTTCTGTCGATGACGACGCGGCGGCGGCGGCGGCGGCGGCGG 480  
 DB 421 CTACCGGCGGCTGCACTTCTGTCGATGACGACGCGGCGGCGGCGGCGGCGGCGGCGG 480  
 QY 481 TCGCTGGAAGACGAGAGCTGATCCGGAAGGCGCGCGGTGTGAGCTGGCGGCGCG 540  
 DB 481 TCGCTGGAAGACGAGAGCTGATCCGGAAGGCGCGCGGTGTGAGCTGGCGGCGCG 540  
 QY 541 GCCTGCGCGGTGCCATCAAGCGCGGCTGCCGAGCATGAAGTGGCGATGCCACACC 600  
 DB 541 GCCTGCGCGGTGCCATCAAGCGCGGCTGCCGAGCATGAAGTGGCGATGCCACACC 600  
 QY 601 AATGCGATGATCCCGGAGATGCGCAATCTCTTCCCTTGTGGAGCTGATGACCTGG 660  
 DB 601 AATGCGATGATCCCGGAGATGCGCAATCTCTTCCCTTGTGGAGCTGATGACCTGG 660  
 QY 661 ACCGTGTTCCAGTGGGGATCAACAACGAGCGCGGCAATCCGGTACCAACCGCATC 720  
 DB 661 ACCGTGTTCCAGTGGGGATCAACAACGAGCGCGGCAATCCGGTACCAACCGCATC 720  
 QY 721 GTGCAATCCGGCGCATCTCTTCCGCTCAACACTTCCCGAGATGTTGGCTACTACAC 780  
 DB 721 GTGCAATCCGGCGCATCTCTTCCGCTCAACACTTCCCGAGATGTTGGCTACTACAC 780

QY 781 GCGCTGAGCGCAGCCTGTTCTTGAGACCATGTCCATGACGCGACGCTTGACATCTGGAG 840  
DB 781 GCGCTGAGCGCAGCCTGTTCTTGAGACCATGTCCATGACGCGACGCTTGACATCTGGAG 840  
QY 841 AAGAACGTGGCCGTGATCGCGCGGCTGAGCTGATCAAGCGCGGCGCGCTGCAAG 900  
DB 841 AAGAACGTGGCCGTGATCGCGCGGCTGAGCTGATCAAGCGCGGCGCGCTGCAAG 900  
QY 901 GACATCGCCATCGAGCTCAACGAGATGTACCGCGAGTGGGACCTTGCTGAATCCGCTCC 960  
DB 901 GACATCGCCATCGAGCTCAACGAGATGTACCGCGAGTGGGACCTTGCTGAATCCGCTCC 960  
QY 961 TTCCGCTATGAGCCACTCTCTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 TTCCGCTATGAGCCACTCTCTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1021 GAGCTGCGCGAGGACATCGACACCGAGCTGAAGCCCGGCTGATGCTGCTGCTGCTGCTG 1080  
DB 1021 GAGCTGCGCGAGGACATCGACACCGAGCTGAAGCCCGGCTGATGCTGCTGCTGCTGCTG 1080  
QY 1081 ATGCTGATGCTGCTGCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 ATGCTGATGCTGCTGCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 ATGCTGCGGAGGAGCGGTCGCGAGACATCAACCGCTTCCGCTGCTGCTGCTGCTGCTG 1200  
DB 1141 ATGCTGCGGAGGAGCGGTCGCGAGACATCAACCGCTTCCGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 ATCATCGGCAC 1212  
DB 1201 ATCATCGGCAC 1212

RESULT 4  
LOCUS E16405 1212 bp DNA linear PAT 28-JUN-1999  
DEFINITION Alcaligenes faecalis gene for creatine amidinohydrolase.  
ACCESSION E16405  
VERSION E16405.1 GI:5711088  
KEYWORDS JP 1998174585-A/1.  
SOURCE Alcaligenes faecalis  
ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
Alcaligenaceae; Alcaligenes.  
REFERENCE 1 (bases 1 to 1212)  
AUTHORS Sogabe,A., Nishiyu,Y. and Kawamura,Y.  
TITLE STABLE CREATINE AMIDINOHYDROLASE  
JOURNAL Patent: JP 1998174585-A 1 30-JUN-1998;  
TOYOBO CO LTD  
OS Alcaligenes faecalis  
PN JP 1998174585-A/1  
PD 30-JUN-1998  
PI 17-DEC-1996 JP 1996337027  
PI SOGABE ATSUSHI, NISHIYA YOSHIKI, KAWAMURA YOSHIOHISA PC  
C12N9/78, C07H21/04, C12N1/21, C12N15/09//C1201/34, C12N9/78, PC  
C12R1.19)  
PC (C12N1/21, C12R1.19);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..1212  
FT /strain='TE3581', /organism='Alcaligenes faecalis' FT  
FT mat\_peptide 1..1212  
FT /product='creatine amidinohydrolase'.  
FEATURES  
source Location/Qualifiers  
1..1212  
/organism='Alcaligenes faecalis'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:511'

ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;  
Best Local Similarity 99.9%; Pred. No. 2.6e-145;  
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGACTGACGACATGTTGACGATGATGAATGACACACGCGCGAAGATTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGATGAATGACACACGCGCGAAGATTATTCGCG 60  
QY 61 TTTTGGATCCGAGATGACCCCGCGCAAAAGACGTTGCGGCTGATGCGCAAGAC 120  
DB 61 TTTTGGATCCGAGATGACCCCGCGCAAAAGACGTTGCGGCTGATGCGCAAGAC 120  
QY 121 AATCTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTATTCGCTGCTG 180  
DB 121 AATCTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTATTCGCTGCTG 180  
QY 181 TACTGCTATTTCCGACGCAAGTACGAGGATGATCGACCAACACGACGACGATT 240  
DB 181 TACTGCTATTTCCGACGCAAGTACGAGGATGATCGACCAACACGACGACGATT 240  
QY 241 TCGGCGGCGATCGACGCGGCGACCTTGCGCGCGAGCTTGGCGCAACATCACTAC 300  
DB 241 TCGGCGGCGATCGACGCGGCGACCTTGCGCGCGAGCTTGGCGCAACATCACTAC 300  
QY 301 ACCGACTGCGGCGCGACATTTCTATCGCGCGTGGCGGACGTCACGCGGCGCAAG 360  
DB 301 ACCGACTGCGGCGCGACATTTCTATCGCGCGTGGCGGACGTCACGCGGCGCAAG 360  
QY 361 CGCATCGGCATCGAGTTCGACACGATCACTCTGACTTCGCGCGCGAGCTCGAGGAAGCC 420  
DB 361 CGCATCGGCATCGAGTTCGACACGATCACTCTGACTTCGCGCGCGAGCTCGAGGAAGCC 420  
QY 421 CTACCGGCGCTGCACTTCTGCAATCGACGACGCTTCATGATGATGCGGACCATTAAG 480  
DB 421 CTACCGGCGCTGCACTTCTGCAATCGACGACGCTTCATGATGATGCGGACCATTAAG 480  
QY 481 TCGCTGCAAGAGAGAGATGATCGCGAAGGCGCGCGCTGCTGCTGCTGCTGCTG 540  
DB 481 TCGCTGCAAGAGAGAGATGATCGCGAAGGCGCGCGCTGCTGCTGCTGCTGCTG 540  
QY 541 GCGTGGCGGCTGCACTCAAGGCGCGGCTGCGGAGCATGAATGCGGATCGCCAGCC 600  
DB 541 GCGTGGCGGCTGCACTCAAGGCGCGGCTGCGGAGCATGAATGCGGATCGCCAGCC 600  
QY 601 AATGCGATGATCCGCGAGATCGCAATCTCTTCCCTTCTGAGGCTGATGACACCTGG 660  
DB 601 AATGCGATGATCCGCGAGATCGCAATCTCTTCCCTTCTGAGGCTGATGACACCTGG 660  
QY 661 ACCTGCTTCAGTCGGGCAATCAACGAGCGGCGGCAATCGGTCACCAACCGGATC 720  
DB 661 ACCTGCTTCAGTCGGGCAATCAACGAGCGGCGGCAATCGGTCACCAACCGGATC 720  
QY 721 GTGCAATCCGCGGACATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACCC 780  
DB 721 GTGCAATCCGCGGACATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACCC 780  
QY 781 GCGCTGAGGCGACGCTGTTCTGCGACATGATGATGAGCGGACGCTGATCTGGAG 840  
DB 781 GCGCTGAGGCGACGCTGTTCTGCGACATGATGATGAGCGGACGCTGATCTGGAG 840  
QY 841 AAGAACGTGGCCGTGATCGCGCGGCTGAGCTGATCAAGCGCGGCGCGCTGCAAG 900  
DB 841 AAGAACGTGGCCGTGATCGCGCGGCTGAGCTGATCAAGCGCGGCGCGCTGCAAG 900  
QY 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTTGCTGAAGTACCGCTCC 960  
DB 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTTGCTGAAGTACCGCTCC 960  
QY 961 TTCCGCTATGAGCCACTCTCTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 TTCCGCTATGAGCCACTCTCTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1021 GAGCTGCGCGAGGACATCGACCGAGCTGAAGCCCGGCGCATGCTGCTCATGAGAGCC 1080

Db 1021 GAGCTGGCGGAGGACATGACACGAGCTGAAGCCCGGACATGCTGATCCATGAGCCG 1080  
Qy 1081 ATGCTGATGCTGCGGAGGAGCATGCGGCTGCGGCGGCTATTCGGAGCACGACATCTTG 1140  
Db 1081 ATGCTGATGCTGCGGAGGAGCATGCGGCTGCGGCGGCTATTCGGAGCACGACATCTTG 1140  
Qy 1141 ATGCTGGGGAGGAGCGGTGCGGAGAACATCACCGGCTTCCGCTTCCGCTCGGAAACACCC 1200  
Db 1141 ATGCTGGGGAGGAGCGGTGCGGAGAACATCACCGGCTTCCGCTTCCGCTCGGAAACACCC 1200  
Qy 1201 ATCATCCGCAAC 1212  
Db 1201 ATCATCCGCAAC 1212

RESULT 5  
BD017699  
LOCUS BD017699 1212 bp DNA linear PAT 27-AUG-2002  
DEFINITION Gene encoding creatine amidinohydrolyase.  
ACCESSION BD017699  
VERSION BD017699.1 GI:22558875  
KEYWORDS JP 2001252088-A/1.  
SOURCE unclassified  
ORGANISM unclassified

REFERENCE  
1 (bases 1 to 1212)  
Sogabe,A., Yamamoto,K. and Kawamura,Y.  
Gene encoding creatine amidinohydrolyase  
Patent: JP 2001252088-A 1 18-SEP-2001;  
TOYOBO CO LTD  
COMMENT OS Alkaligenes faecalis TE3581 (FERM P14237)  
PN JP 2001252088-A/1  
PD 18-SEP-2001  
PF 26-FEB-2001 JP 2001051054  
PI ATSUSHI SOGABE,KAZUMI YAMAMOTO,YOSHIHISA KAWAMURA PC  
C12N15/09,C12N1/21,C12N9/78//C12N15/09,C12R1.425/,C12N1/21, PC  
C12R1.425/,  
PC (C12N9/78,C12R1:05),C12N15/00,(C12N15/00,C12R1:425) CC Gene  
encoding creatine amidinohydrolyase  
FH Key Location/Qualifiers  
FT source 1..1212  
FT 1..1212 Location/Qualifiers  
FT P14237'.  
Location/Qualifiers  
1..1212 /organism="Alkaligenes faecalis TE3581 (FERM  
P14237)"  
source /organism="unclassified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;  
Best Local Similarity 99.9%; Pred. No. 2.6e-145;  
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGACTACGACATGTTGCACTGATGAAATGCGACAAAGCGCGAAGATTTATTCGCCG 60  
Db 1 ATGACTACGACATGTTGCACTGATGAAATGCGACAAAGCGCGAAGATTTATTCGCCG 60  
Qy 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGAGCGTGGCGGCTGGATGGCCAAAGC 120  
Db 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGAGCGTGGCGGCTGGATGGCCAAAGC 120  
Qy 121 AATGTCGATGCGGCGCTGTTCACTCTTATCAGTCACTCACTAATTCGCGCTGCTG 180  
Db 121 AATGTCGATGCGGCGCTGTTCACTCTTATCAGTCACTCACTAATTCGCGCTGCTG 180  
Qy 181 TACTGCTATTTCGAGCGCAAGTACGGCATGCTCATCGACCAACAGCGCAGCATT 240  
Db 181 TACTGCTATTTCGAGCGCAAGTACGGCATGCTCATCGACCAACAGCGCAGCATT 240  
Qy 241 TCGGCGGCGCATCGACGGCGGCGACGCTGGCGCGGCGGCTTCGCGGCAATCATCACTAC 300  
Db 241 TCGGCGGCGCATCGACGGCGGCGACGCTGGCGCGGCGGCTTCGCGGCAATCATCACTAC 300

Db 241 TCGGCGGCGCATCGACGGCGGCGACGCTGGCGCGGCGGCTTCGCGGCAATCATCACTAC 300  
Qy 301 ACCGACTGGCGCGCGCGCAATTTCTATGCGCGCTGCGCACTGACCAAGCGCGCAAG 360  
Db 301 ACCGACTGGCGCGCGCGCAATTTCTATGCGCGCTGCGCGCACTGACCAAGCGCGCAAG 360  
Qy 361 CGCATTCGCGCATTCGAGTTCGACACGTCGATTCGACCTTCGCGCGCGCGCTCGAGAAAGCC 420  
Db 361 CGCATTCGCGCATTCGAGTTCGACACGTCGATTCGACCTTCGCGCGCGCGCTCGAGAAAGCC 420  
Qy 421 CTACCGGCGCTGCACTTCGTCGACATCAGCGCAGCTTCGATGTCGTCGCGCAGCATCAAG 480  
Db 421 CTACCGGCGCTGCACTTCGTCGACATCAGCGCAGCTTCGATGTCGTCGCGCAGCATCAAG 480  
Qy 481 TCGCTCGAAGAGCAAGAGCTGATTCGCGAAGCGCGCGCTGTCGATGTCGCGCGCGG 540  
Db 481 TCGCTCGAAGAGCAAGAGCTGATTCGCGAAGCGCGCGCTGTCGATGTCGCGCGCGG 540  
Qy 541 GCGTGGCGGCTGCGCATCGACGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
Db 541 GCGTGGCGGCTGCGCATCGACGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
Qy 601 AATGCGATGATTCGCGGAGATCGCCAAATCGTTCCCTTCGTCGAGCTGATGSAACCTGG 660  
Db 601 AATGCGATGATTCGCGGAGATCGCCAAATCGTTCCCTTCGTCGAGCTGATGSAACCTGG 660  
Qy 661 ACCGTGTTCCAGTTCGCGGATCAACACCGACGCGCGCGCAATTCGCTACCAACCGCATC 720  
Db 661 ACCGTGTTCCAGTTCGCGGATCAACACCGACGCGCGCGCAATTCGCTACCAACCGCATC 720  
Qy 721 GTGCAATTCGCGGCAATTCCTTTCGTCGCAACCTTCGCGGATGATTCGCTGCTACTAAC 780  
Db 721 GTGCAATTCGCGGCAATTCCTTTCGTCGCAACCTTCGCGGATGATTCGCTGCTACTAAC 780  
Qy 781 GCGCTGAGCGGCAAGCTGCTTCGCGACCATGTCGATGATGCGCGCAGCTTCGATGCGAG 840  
Db 781 GCGCTGAGCGGCAAGCTGCTTCGCGACCATGTCGATGATGCGCGCAGCTTCGATGCGAG 840  
Qy 841 AAGAACGTGGCGCTGTCATGCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900  
Db 841 AAGAACGTGGCGCTGTCATGCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900  
Qy 901 GAGATGCGCATTCGAGCTCAACGAGATGTAACCGGAGTGGGACCTGCTGAGTACCGCTCC 960  
Db 901 GAGATGCGCATTCGAGCTCAACGAGATGTAACCGGAGTGGGACCTGCTGAGTACCGCTCC 960  
Qy 961 TTGCGGATAGGCACTCTTCGCGGCTGCTGTCGCACTACTACGCTTCGCGAGCGCGGCTG 1020  
Db 961 TTGCGGATAGGCACTCTTCGCGGCTGCTGTCGCACTACTACGCTTCGCGAGCGCGGCTG 1020  
Qy 1021 GAGCTGGCGGAGCATTCGACACCGAGCTGAGAGCTGGAGCTTCGATGAGAGCCG 1080  
Db 1021 GAGCTGGCGGAGCATTCGACACCGAGCTGAGAGCTGGAGCTTCGATGAGAGCCG 1080  
Qy 1081 AATGTCATGTCGCGGAGGAGCATGCGCGGCGCGGCGGCTGATGCGGAGCAATCTG 1140  
Db 1081 AATGTCATGTCGCGGAGGAGCATGCGCGGCGCGGCGGCTGATGCGGAGCAATCTG 1140  
Qy 1141 ATGCTGGGGAGGAGCGGTGCGGAGAACATCAACCGGCTTCGCTTCGCTCGGAAACACCC 1200  
Db 1141 ATGCTGGGGAGGAGCGGTGCGGAGAACATCAACCGGCTTCGCTTCGCTCGGAAACACCC 1200  
Qy 1201 ATCATCCGCAAC 1212  
Db 1201 ATCATCCGCAAC 1212

RESULT 6  
BD105662  
LOCUS BD105662 1212 bp DNA linear PAT 27-AUG-2002  
DEFINITION Stable creatine amidinohydrolyase.  
ACCESSION BD105662  
VERSION BD105662.1 GI:22651236

KEYWORDS JP 2001346594-A/1.  
SOURCE unidentified  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1212)  
AUTHORS Sogabe,A., Mishiya,Y. and Kawamura,Y.  
TITLE Stable creatine amidinohydrolase  
JOURNAL Patent: JP 2001346594-A 1 18-DEC-2001;  
TOYOBO CO LTD  
OS Alcaligenes faecalis  
PN JP 2001346594-A/1  
PD 18-DEC-2001  
PF 19-APR-2001 JP 2001121708  
PI ATSUSHI SOGABE, YOSHIKAZU NISHIYAMA, YOSHIOHISA KAWAMURA, PC  
C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/80// PC  
C12N9/80, C12N1/05, C12N15/00, C12N5/00  
CC Strandedness: Double;  
CC Topology: Linear;  
CC Stable creatine amidinohydrolase  
FH key Location/Qualifiers  
FT source 1..1212  
Location/Qualifiers  
1..1212 /organism='Alcaligenes faecalis'.  
source  
FEATURES  
source 1..1212  
Location/Qualifiers  
1..1212 /organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'  
ORIGIN  
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;  
Best Local Similarity 99.9%; Pred. No. 2.6e-145;  
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGACTGACGACATGTTGACGATGAAATGACACACGCGAGAAAGATTATTGCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGAAATGACACACGCGAGAAAGATTATTGCGCG 60  
QY 61 TTTTGGATGCGGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120  
DB 61 TTTTGGATGCGGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120  
QY 121 AATGTCAGTGGCGCGCTGTTCACTCTTATCACTGATCACTAATTTCCGGCTGGCTG 180  
DB 121 AATGTCAGTGGCGCGCTGTTCACTCTTATCACTGATCACTAATTTCCGGCTGGCTG 180  
QY 181 TACTGCTATTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 240  
DB 181 TACTGCTATTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 240  
QY 241 TCGGCGCGGATGACG 300  
DB 241 TCGGCGCGGATGACG 300  
QY 241 TCGGCGCGGATGACG 300  
DB 241 TCGGCGCGGATGACG 300  
QY 301 ACCGACTGGCG 360  
DB 301 ACCGACTGGCG 360  
QY 301 ACCGACTGGCG 360  
DB 301 ACCGACTGGCG 360  
QY 361 CGCATGCGGATGAGTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 420  
DB 361 CGCATGCGGATGAGTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 420  
QY 421 CTACCGGCGGATGAGTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 480  
DB 421 CTACCGGCGGATGAGTTGACGACGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 480  
QY 481 TCGCTGAG 540  
DB 481 TCGCTGAG 540  
QY 541 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
DB 541 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600

QY 601 AATGCGATGATCCGCGAGATGCGCAATCGTTCCCTTGGTGGAGCTGATGACACCTGG 660  
DB 601 AATGCGATGATCCGCGAGATGCGCAATCGTTCCCTTGGTGGAGCTGATGACACCTGG 660  
QY 661 ACCGTTTCCAGTGGCGGATGACACGAGCGCGCGGACATCCGCTGACCAACCGCATC 720  
DB 661 ACCGTTTCCAGTGGCGGATGACACGAGCGCGCGGACATCCGCTGACCAACCGCATC 720  
QY 721 GTGCAATCCGCGGATGATCTTTGCTGATCACTCTCCGATGATCTTCCGCTACTACAC 780  
DB 721 GTGCAATCCGCGGATGATCTTTGCTGATCACTCTCCGATGATCTTCCGCTACTACAC 780  
QY 781 GCGTGGAGCGGACGATGTTCTGCGACGATGCGATGACGCGGACCTGATCATCTGGAG 840  
DB 781 GCGTGGAGCGGACGATGTTCTGCGACGATGCGATGACGCGGACCTGATCATCTGGAG 840  
QY 841 AAGAACGTGGCGGATGATCGCGCGGCGGCTGAGCTGATCAAGCGGCGCGGCTGCAAG 900  
DB 841 AAGAACGTGGCGGATGATCGCGCGGCGGCTGAGCTGATCAAGCGGCGCGGCTGCAAG 900  
QY 901 GACATCGGCGGATGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
DB 901 GACATCGGCGGATGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
QY 961 TTGGGCTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020  
DB 961 TTGGGCTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020  
QY 1021 GAGCTGGCGGCGGAGATGACACCGAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
DB 1021 GAGCTGGCGGCGGAGATGACACCGAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
QY 1081 ATGCTGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
DB 1081 ATGCTGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
QY 1141 ATGCTGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200  
DB 1141 ATGCTGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200  
QY 1201 ATCATCCGCAAC 1212  
DB 1201 ATCATCCGCAAC 1212  
RESULT 7  
AB016788 1215 bp DNA linear BCT 06-JAN-2000  
LOCUS  
DEFINITION  
AB016788 1215 bp DNA linear BCT 06-JAN-2000  
strain:KS-85.  
ACCESSION  
AB016788.1 GI:6681665  
VERSION  
AB016788.1 GI:6681665  
KEYWORDS  
Alcaligenes sp.  
Alcaligenes sp.  
Alcaligenes sp.  
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
Alcaligenaceae; Alcaligenes.  
REFERENCE  
1 (sites)  
Furukawa,K., Ichikawa,T., Koyama,Y. and Suzuki,M.  
AUTHORS  
Molecular cloning and sequence analysis of the gene encoding  
thermostable creatinase from Alcaligenes sp. KS-85  
TITLE  
Unpublished  
JOURNAL  
2 (bases 1 to 1215)  
Koyama,Y. and Furukawa,K.  
REFERENCE  
Direct Submission  
Submitted (05-AUG-1998) Yasuji Koyama, Kikkoman Corporation,  
Research and Development, Noda 399, Noda, Chiba 278-0037, Japan  
(E-mail:daiz-kk@noda399-net.ne.jp, Tel:+81-471-23-5571,  
Fax:+81-471-23-5959)  
FEATURES  
source 1..1215  
Location/Qualifiers  
1..1215 /organism='Alcaligenes sp.'  
/mol\_type='genomic DNA'

## CDS

```
/strain="KS-85"  
/db_xref="taxon:512"  
1..1215  
/codon_start=1  
/transl_table=11  
/product="creatine amidinohydrolase"  
/protein_id="BA88830.1"  
/db_xref="GI:6681666"  
/translation="MTDDMLHVMKENGEDKDYSPSDAEMTRONDYRGMAKNVDA  
ALFTSYHCINYSGLYCFFGRKXGMYIDHNNAITTSAGIDGGPMRSPEDNITTD  
WRDNFRAVROLTTGAKRIGIEFDHNLDFRQLEALPEVEFMDISOSMWRITK  
SLEBOKLIREGARVCDVGAACAATKAGVEHEVALATTNMAIRLAKSPFVELND  
TMTFOSGINTDGAHNPVTNRIVOSGDIISLNTPEPMI RGYTALERTLDFHDNDAL  
DIWKAVAVHRRGLIELKPGARCDIIELEMTREMDLKYRSPFGHSGEVICHTY  
GREXGVELREDIDTELKPGMVSMPEVMPLPEGMPAGGYREHDILIVGEDAENITG  
PPFGEHNHIRN"
```

## ORIGIN

```
Query Match 99.5%; Score 1205.6; DB 1; Length 1215;  
Best Local Similarity 99.7%; Pred. No. 1,1e-144;  
Matches 1208; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 ATGACTGACGACATGTTGACGATGATAATGACAAACGCGAGAAAGATTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGACGATGATAATGACAAACGCGAGAAAGATTATTCGCG 60  
  
QY 61 TTTTCGAGATGCGAGATGACCCGCGCCCAAAACGACGTTGCGGCTGGATGCCAAGAC 120  
DB 61 TTTTCGAGATGCGAGATGACCCGCGCCCAAAACGACGTTGCGGCTGGATGCCAAGAC 120  
  
QY 121 AATGTCGATGCGGCGCTGTTTACCTTTATCACTGACATCACTAATTCGCGGCTG 180  
DB 121 AATGTCGATGCGGCGCTGTTTACCTTTATCACTGACATCACTAATTCGCGGCTG 180  
  
QY 181 TACTGCTATTTGCGACGCAAGTACGCGATGCTATCGACCAACACGCGACGACGATT 240  
DB 181 TACTGCTATTTGCGACGCAAGTACGCGATGCTATCGACCAACACGCGACGACGATT 240  
  
QY 241 TCGGCGCGGATCGACGCGCGGCGACGCTGCGCGCGACGTTGGCGACACATCACTAC 300  
DB 241 TCGGCGCGGATCGACGCGCGGCGACGCTGCGCGCGACGTTGGCGACACATCACTAC 300  
  
QY 301 ACCGACTGCGCGCGCGCAATTTGTATCGCGCGCGCGACGACCAACGCGCGCGCAAG 360  
DB 301 ACCGACTGCGCGCGCGCAATTTGTATCGCGCGCGCGACGACCAACGCGCGCGCAAG 360  
  
QY 361 CGCATCGGACATCGAGTTGACACGTCATCTTCCGCGCGCGACGCTGAGAGAGCC 420  
DB 361 CGCATCGGACATCGAGTTGACACGTCATCTTCCGCGCGCGACGCTGAGAGAGCC 420  
  
QY 421 CTACCGGCGGTCGACTTGTGACATGACCGACGCTTGTGATGATGCGACCATCAAG 480  
DB 421 CTACCGGCGGTCGACTTGTGACATGACCGACGCTTGTGATGATGCGACCATCAAG 480  
  
QY 481 TCGCTCGAAGAGCAAGCTGATCCGCGAAGCGCGCGCGTGTGACGTCGCGCGCGCG 540  
DB 481 TCGCTCGAAGAGCAAGCTGATCCGCGAAGCGCGCGCGTGTGACGTCGCGCGCGCG 540  
  
QY 541 GCGTGCAGGCTGACATCAAGCGCGCGCGCGCGAGCATGAAGTGGCGATGCCACACC 600  
DB 541 GCGTGCAGGCTGACATCAAGCGCGCGCGCGCGAGCATGAAGTGGCGATGCCACACC 600  
  
QY 601 AATGCGATGATCGCGAGATGCGCAAAATCGTTCCCTTCGCGAGCTGATGACACCTGG 660  
DB 601 AATGCGATGATCGCGAGATGCGCAAAATCGTTCCCTTCGCGAGCTGATGACACCTGG 660  
  
QY 661 ACCGCTTTCAGTGGCGCATCAACGCGCGCGCGACATTCGGTCAACACCGCATC 720  
DB 661 ACCGCTTTCAGTGGCGCATCAACGCGCGCGCGACATTCGGTCAACACCGCATC 720  
  
QY 721 GTGCAATCGGCGCATCTTTCGCTCAACACCTTCGCGATGATCTTCGCTACTACAC 780  
DB 721 GTGCAATCGGCGCATCTTTCGCTCAACACCTTCGCGATGATCTTCGCTACTACAC 780
```

```
QY 781 GCGCTGAGAGGACGCTGTTCTGCGACCATGTGATACCGCAGCCTGACATCTGGAG 840  
DB 781 GCGCTGAGAGGACGCTGTTCTGCGACCATGTGATACCGCAGCCTGACATCTGGAG 840  
  
QY 841 AAGAACGTGCGCGTGCATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAG 900  
DB 841 AAGAACGTGCGCGTGCATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAG 900  
  
QY 901 GACATCGCATTCGAGCTCAACGAGATGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960  
DB 901 GACATCGCATTCGAGCTCAACGAGATGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960  
  
QY 961 TTGCGCTATGCGCCACTCTCTCGCGCTGCTGTGCACTACAGGTCCGCGAGCGCGCTG 1020  
DB 961 TTGCGCTATGCGCCACTCTCTCGCGCTGCTGTGCACTACAGGTCCGCGAGCGCGCTG 1020  
  
QY 1021 GAGCTGCGAGGACATCGACACCGAGCTGAAGCGCGCATGCGTCTCCATGAGCGG 1080  
DB 1021 GAGCTGCGAGGACATCGACACCGAGCTGAAGCGCGCATGCGTCTCCATGAGCGG 1080  
  
QY 1081 ATGCTGATGCTGCGCGAGGCGCATGCGCGTGCCTATGCGAGCAGACATCTCTG 1140  
DB 1081 ATGCTGATGCTGCGCGAGGCGCATGCGCGTGCCTATGCGAGCAGACATCTCTG 1140  
  
QY 1141 ATGCTGCGGAGGACGCTGCGGAGAACATACACGCGCTTCCGCTGCGTCCGGAACAGCC 1200  
DB 1141 ATGCTGCGGAGGACGCTGCGGAGAACATACACGCGCTTCCGCTGCGTCCGGAACAGCC 1200  
  
QY 1201 ATCATCCGCAAC 1212  
DB 1201 ATCATCCGCAAC 1212
```

```
RESULT 8  
E11155 1215 bp DNA linear PAT 29-SEP-1997  
LOCUS E11155  
DEFINITION gDNA encoding creatine amidinohydrolase.  
ACCESSION E11155  
VERSION E1115.1 GI:22024796  
KEYWORDS JP 1996089255-A/1.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1215)  
AUTHORS Furukawa,K., Ichikawa,T., Suzuki,M. and Koyama,T.  
TITLE NOVEL, CREATINE AMIDINOHYDROLASE GENE, NOVEL, RECOMBINANT DNA AND  
PRODUCTION OF CREATINE AMIDINOHYDROLASE  
JOURNAL Patent: JP 1996089255-A 1 09-Apr-1996;  
KIKKOMAN CORP  
COMMENT  
OS Alca1genes sp. KS-85  
PN JP 1996089255-A/1  
PD 09-APR-1996  
PF 29-SEP-1994 JP 1994235737  
PI FURUKAWA KEISUKE, ICHIKAWA TOSHIO, SUZUKI MASARU, KOYAMA TAIJI  
PC C12N15/09,C12N9/78,(C12N9/78,C12R1:05),(C12N9/78,C12R1:19); CC  
strandedness: Double;  
CC topology: Linear;  
FH Key  
FH Location/Qualifiers  
FT source 1..1215 /organism='Alca1genes sp. KS-85' FT CDS  
FT 1..1215 /product='Creatine amidinohydrolase'.
```

## FEATURES

```
source  
1..1215  
/location/Qualifiers  
/organism="unidentified"  
/mol_type="genomic DNA"  
/db_xref="taxon:32644"
```

## ORIGIN

```
Query Match 97.8%; Score 1184.8; DB 6; Length 1215;  
Best Local Similarity 98.6%; Pred. No. 4.8e-142;
```

Matches 1195; Conservative 0; Mismatches 17; Indels 0; Gaps 0;				
QY	1	ATGACTGAGACATGTTTGCAGTGAATGAGCAAAAGGCGAAGAAATTAATTCGCG	60	
DB	1	ATGACTGAGACATGTTTGCAGTGAATGAGCAAAAGGCGAAGAAATTAATTCGCG	60	
QY	61	TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGGATGGCGAAGAC	120	
DB	61	TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGGATGGCGAAGAC	120	
QY	121	AATGTGATGCGGCGGTGTACCTTTATCATCTGATCAACTATTCCGGTGGCTG	180	
DB	121	AATGTGATGCGGCGGTGTACCTTTATCATCTGATCAACTATTCCGGTGGCTG	180	
QY	181	TACTGCTATTTTGGAGAGCAAGTACGGCATGTGATGACCAAAAGCGCACAGAT	240	
DB	181	TACTGCTATTTTGGAGAGCAAGTACGGCATGTGATGACCAAAAGCGCACAGAT	240	
QY	241	TGGCGCGGATGACGCGCGCGCAAGCTTGGCGCGGAGCTTGGCGAACAATCACTAC	300	
DB	241	TGGCGCGGATGACGCGCGCGCAAGCTTGGCGCGGAGCTTGGCGAACAATCACTAC	300	
QY	301	ACCGACTGCGCGCGCAATTTTATTCGCGCGGTGCGCAGCTGACCAAGGCGCGCAAG	360	
DB	301	ACCGACTGCGCGCGCAATTTTATTCGCGCGGTGCGCAGCTGACCAAGGCGCGCAAG	360	
QY	361	CGCATGGCATGCAATTGCAACGTCATTTGATCTTCCGCGCGCACTCGAAGAGCC	420	
DB	361	CGCATGGCATGCAATTGCAACGTCATTTGATCTTCCGCGCGCACTCGAAGAGAGC	420	
QY	421	CTAACGGGGGTGCACTTGGTGCATCAGCAGCGCTGATGATGATGGGCAATCAAG	480	
DB	421	CTAACGGGGGTGCACTTGGTGCATCAGCAGCGCTGATGATGATGGGCAATCAAG	480	
QY	481	TCGCTCGAAGAGAGCAAGTGTATCCGGAAGCGCGCGGTGTGATCGCGCGCG	540	
DB	481	TCGCTCGAAGAGAGCAAGTGTATCCGGAAGCGCGCGGTGTGATCGCGCGCG	540	
QY	541	GCGTGGCGGCTGCCATCAAAGCGCGCGGTGCGCGAGCATGAGTGGCATCCGACAC	600	
DB	541	GCGTGGCGGCTGCCATCAAAGCGCGCGGTGCGCGAGCATGAGTGGCATCCGACAC	600	
QY	601	AATGCGATGATCGCGAGATCGCAATGCTTCCCTTGTGGAGCTGATGAGACCTGG	660	
DB	601	AATGCGATGATCGCGAGATCGCAATGCTTCCCTTGTGGAGCTGATGAGACCTGG	660	
QY	661	ACCTGCTTCAGTCCGGCATCAACACGCGCGCGCAATCCGCTCAACCAACGCGATC	720	
DB	661	ACCTGCTTCAGTCCGGCATCAACACGCGCGCGCAATCCGCTCAACCAACGCGATC	720	
QY	721	GTGCAATCCGGCGACATCTTTGCTCAACACTTCCGATGATCTTGGCTACTACAC	780	
DB	721	GTGCAATCCGGCGACATCTTTGCTCAACACTTCCGATGATCTTGGCTACTACAC	780	
QY	781	GCGCTGGAGCGGAGCTGTTTGGCGACATGTCGATGAGAGCGGAGCTTGGAGAG	840	
DB	781	GCGCTGGAGCGGAGCTGTTTGGCGACATGTCGATGAGAGCGGAGCTTGGAGAG	840	
QY	841	AAGAAGTGGCGGTGATCGCGCGCGGTGAGCTGATCAAGCGCGCGCGCTGCAAG	900	
DB	841	AAGAAGTGGCGGTGATCGCGCGCGGTGAGCTGATCAAGCGCGCGCGCTGCAAG	900	
QY	901	GACATGCGCATGAGCTCAACAGAGATGTAACCGGATGAGGAGCTGCTGAAGTAC	960	
DB	901	GACATGCGCATGAGCTCAACAGAGATGTAACCGGATGAGGAGCTGCTGAAGTAC	960	
QY	961	TTTGGCTATGGCACTCTTGGCGGTGCTGCACTAATGATGAGGCGCGCGCTG	1020	
DB	961	TTTGGCTATGGCACTCTTGGCGGTGCTGCACTAATGATGAGGCGCGCGCTG	1020	
QY	1021	GAGCTGCGGAGAGCATGCAACCGAGCTGAAGCGCGCATGTGTCTTCATGAGAGCG	1080	
DB	1021	GAGCTGCGGAGAGCATGCAACCGAGCTGAAGCGCGCATGTGTCTTCATGAGAGCG	1080	

QY	1081	ATGCTGATGCTGCCGAGAGGATGCGCGTCCGCGGCTATCCGAGACACATCCGTG	1140	
DB	1081	ATGCTGATGCTGCCGAGAGGATGCGCGTCCGCGGCTATCCGAGACACATCCGTG	1140	
QY	1141	ATCGTGGGAGAGAGCGTGCAGAGAACATACCGGCTTCCGTTGCTCCGGAACACAC	1200	
DB	1141	ATCGTGGGAGAGAGCGTGCAGAGAACATACCGGCTTCCGTTGCTCCGGAACACAC	1200	
QY	1201	ATCATCCGCAAC 1212		
DB	1201	ATCATCCGCAAC 1212		
RESULT 9				
AX721933				
LOCUS				
DEFINITION				
Sequence 1 from Patent EP1298213.				
ACCESSION				
AX721933				
VERSION				
AX721933.1 GI:30422515				
KEYWORDS				
SOURCE				
ORGANISM				
Erwinia sp.				
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
Enterobacteriaceae; Erwinia.				
REFERENCE				
1 Shao, Z., Schmuck, R., Kratzsch, P., Kenkies, J. and Weisner, H.				
AUTHORS				
TITLE				
Variants of an erwinia-type creatinase				
JOURNAL				
Patent: EP 1298213-A 1 02-APR-2003;				
Roche Diagnostics GmbH (DE) / F. HOFMANN-LA ROCHE AG (CH)				
FEATURES				
source				
1..1215				
/organism="Erwinia sp."				
/mol_type="unassigned DNA"				
/db_xref="taxon:558"				
1..1215				
/note="unnamed protein product"				
/codon_start=1				
/transl_table=1				
/protein_id="CAD90216.1"				
/db_xref="GI:30422516"				
/translation="MTDMLHVMKHNKEKEYSPSPDAEMTRROSDVRRMAENDVDA				
ALPTSYHCINYYSGELVCEYKRGKMYIDODHATTISAGIGOPWRSSFGDNTTYTD				
WRPNFQAVNQLTPRGARIGIEPDHNPRLTEFLPEVEFDIGQPMWMTYK				
SLEBOKLIRBGRICIDVGGAQVAAYAGVEHEVAZATTAMREIKSPPEYEMD				
TWTFQSGINTDGAHPVYNIIVSGDILSNTPMLFGYTTALERTLPCDHVDASL				
DIWEKNVAVHRGIELIKPGARCKDIIIELEMYREWDLYRSFGYGHSPGVLSHY				
GREAGVELREDIDTVLPQMVVSMPEWMLPEGAPGAGYREHDLIVGEDGAENITG				
FPPEPHNIEN"				
ORIGIN				
Query Match				
Best Local Similarity 81.0%; Score 981.6; DB 6; Length 1215;				
Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;				
QY	1	ATGACTGAGACATGTTTGCAGTGAATGAGCAAAAGGCGAAGAAATTAATTCGCG	60	
DB	1	ATGACTGAGACATGTTTGCAGTGAATGAGCAAAAGGCGAAGAAATTAATTCGCG	60	
QY	61	TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGGATGGCGAAGAC	120	
DB	61	TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGGATGGCGAAGAC	120	
QY	121	AATGTGATGCGGCGGTGTACCTTTATCATCTGATCAACTATTCCGGTGGCTG	180	
DB	121	AATGTGATGCGGCGGTGTACCTTTATCATCTGATCAACTATTCCGGTGGCTG	180	
QY	181	TACTGCTATTTTGGAGAGCAAGTACGGCATGTGATGACCAAAAGCGCACAGAT	240	
DB	181	TACTGCTATTTTGGAGAGCAAGTACGGCATGTGATGACCAAAAGCGCACAGAT	240	
QY	241	TGGCGCGGATGACGCGCGCGCAAGCTTGGCGCGGAGCTTGGCGAACAATCACTAC	300	
DB	241	TGGCGCGGATGACGCGCGCGCAAGCTTGGCGCGGAGCTTGGCGAACAATCACTAC	300	



Dp	241	TCGGCCGGCATGATGGCGGTGAGCCCTGGCCGCGGTAGCTTGGCGCAACAATCACTAAT	300
Qy	301	ACCGAATGGCGCCCGCACAATTTCTATTCGGCGCTGTGGCCAGCTGACCAACGGCGCCAAAG	360
Dp	301	ACGGATTTGGCGCCGCGACAACTTCTACACAGCGCGTCCGCCAATCAACCCCGCGCCAGG	360
Qy	361	CGCATCGGATCGAGTTTCGACCAACGATATCTCGATTCGCGCGCCAGGCTCGAGGAAGCC	420
Dp	361	CGCATCGGATCGAGTTTCGATCACGTAACCTTGACTTTCGCGCCACGCTTCGAAAGAGCG	420
Qy	421	CTACCGGGGCGTGACTTCTGTGACATCAACCAACCCCTCGATGTGAATGCGCAATCAAG	480
Dp	421	CTGCCCGGCGTGAAGTTCTGTGATTCGGTCAACCGTTCATGTGATGTGCGAACGCTCAAG	480
Qy	481	TGGCTTCGAAGACAGAAGCTGATTCGCGGAAGCGCCCGGTGTGTACGTTCGCGCGCGG	540
Dp	481	TGGCTTCGAAGACAGAAGCTGATTCGCGAGGGTCCCGATCTGCAACGTTCGCGCGGTGCC	540
Qy	541	GCGTCGCGCGCTGSCATCAAGGCCGGCGTACCAGCATGAAGTGGCGATTCGCACACC	600
Dp	541	GCGTCGCTTGGCCGCGCTCAAGGCCGCGCTTCCGAGACAAGAGTCCGATTCGCACAAAC	600
Qy	601	AATGCGATGATCCGCGAGATCGCCAAATCGTTCCCTTCGTGGAGCTGATGACAACCTGG	660
Dp	601	AATGCGATGATCCGCGAGATCGCAAGATCGCAAGTGTTCCTTCGTGGAATGATGACACCTGG	660
Qy	661	ACCTGGTTTCAGTCCGGCATCAACAACGAGCGCGCCACAAATCCGCTACCAACCGCATC	720
Dp	661	ACCTGGTTTCAGTCCGGCATCAACAACGAGCGGCCCAAAATCCGCTGACCAACCGCATC	720
Qy	721	GTGCAATCCGCGGAGCATCTTTCGCTCAACACTTCCGATGATCTTCGGCTAATCAACG	780
Dp	721	GTGCAATCCGCGGAGATATCTCTGTGCTCAACAGTTCCGATATCTTCGGCTAATCAACG	780
Qy	781	GCGCTGAGCGCAGCTGTCTTCGACCAATGTGATGACGCCAGCCTCGACATCTTGGAG	840
Dp	781	GCGCTGAGCGAACCCTGTTCTGCAACCACTTCGACGATGCGACGCTTGAATCTTGGAG	840
Qy	841	AAGAACGTGGCGGTGATCACCAGCGCGGCTCGAGCTGATCAACCGCGGCGCGGCTGCAAG	900
Dp	841	AAGAACGTGGCGGTGATCACCAGCGCGGCTCGGAATCTATCAACCGCGGTGCGGCTGCAAG	900
Qy	901	GACATCGGCATGAGCTCAACAGATGTACCCGAGTGGGACCTGCTGAATACGCTCC	960
Dp	901	GATATCGGCATGAACTCAACGAGATGTACCGGAGTGGGACTCTGTAATGATCCGCTCC	960
Qy	961	TTCCGCTATGCGCACTCTTCGGCGGTGTGTGCACTACTAGTGTGCGAGGCGGCGGTG	1020
Dp	961	TTCCGCTATCGGCACTCTTCGGCGGTGTCTCCACTACTAGCGCGGAGGCGGCGGTG	1020
Qy	1021	GAGCTGCCCGAGAGATATGACACCGAGCTGAAAGCCCGGATATGTGTGTCTTCATGAGCCG	1080
Dp	1021	GAACTGCCCGAGAGATATGATACGTGTGACAGCCCGGATATGTGTGTCTTCATATGAGCCG	1080
Qy	1081	ATGTGATGTGTCGCGAGGAGGAGTGCCTCGGATGCGCGGCTATCGAGAGACGACATCTCG	1140
Dp	1081	ATGTGATGTGTCGAGAGAGGCGCTCCGCGCGGCGGCTACCGGAGACGACATCTCTG	1140
Qy	1141	ATCGTCGGGGAGAGCGGTGCGGAATCAACCGGCTTCCGTTGGTTCGGACAAC	1200
Dp	1141	ATCGTCGGGGAGAGCGGCGAGAGAATTAACGGATTCCTCTTCGGGCTGAGACAAC	1200
Qy	1201	ATCATTCGCAAC	1212
Dp	1201	ATCATTCGCAAC	1212

RESULT 10					
AX721949	AX721949	1215 bp	DNA	linear	PAT 07-MAY-2003
LOCUS	Sequence	17	from Patent EP1298213.		
DEFINITION	AX721949				
ACCESSION	AX721949.1	GI:30422525			
VERSION					

	KEYWORDS	. synthetic construct
	SOURCE	synthetic construct
	ORGANISM	artificial sequences.
	REFERENCE	1 Shao Z., Schmuck R., Kratzsch P., Kenkies J. and Weisner H. Variants of an erwinia-type Creelinaase Patent: EP 1298213-A 17 02-AR-2003; Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
	AUTHORS	JOURNAL
	TITLE	location/Qualifiers
	FEATURES	1..1215
	source	/organism="synthetic construct"
	CDS	/mol_type="unassigned DNA" /db_xref="taxon:32630" /note="variant Ctgcc2" 1..1215 /note="unnamed protein product" /codon_start=1 /transl_table=1 /protein_id="CAD90221.1" /db_xref="GI:30422526" translation:"MTDDMLVHKMHNGEKEYSFSDAEWTRHROSDVRKRWALENDYDA ALTSTGYCNINVSGLVCYGRGKYGVIVODDATTTLSAGIDGGQPPRRSGDNIITVD WRDNRYQAIRGRLTPBARRIGIFPDHYDLPRFTTBELPLGVSFVIDIQQSSMMRTTYK SLSEQLIKREGARI CDVGGAACVAAYAKGVPEHEVALATTNAVNRILASFPVEELD TWMFGINTGDANHNVNRIIVOSDIISLNTFPMI PGYTVALSETLPEDHDADSL DTWEKVAVHRGLELI KEGARCKDI AIELTENYREMDLIKYS FGCGHSFYGLSHSY GRAGVELREDIDPTVL QPGMVMSMEPWML PECAFGAGYREHIDL I VGEDAENITG FFPGEPHNIIRN"
ORIGIN	Query Match	81.0%; Score 981.6; DB 6; Length 1215;
	Best Local Similarity	88.1%; Pred. No. 3.9e-116;
	Matches 1068; Conservative	0; Mismatches 144; Indels 0; Gaps 0;
Oy	1 ATGACTGCACGACATGTTCACCTTGATGAATAATGGCACACA CGGCGAGAAGATTATTGGCCG 60	
Dd	1 ATGACTGCACGACATGTTCACCTTGATGAATAATGGCACACA ATGATGTAAGAAATATTCGCCC 60	
Oy	61 TTTTCGGATCCGACATGTGACCGCGCCCAAAGAAGCTTCGCGCGGTGATGGCCAGAAGAAC 120	
Dd	61 TTTTCGGATCCGACATGTGACCGCGCCCAAAGAGTAGCTGGCGCGCTGATGCGCGAAGAAC 120	
Oy	121 AATGTGATGCGCGCGCTGTTCACCTTTATCATCTGATCAATACTATTCCGCTGGCTG 180	
Dd	121 GAGGTGAGCGCTGCGGTTCACCTCTATCATTTGATGATCACTAATACTAATCTTGATGGCTG 180	
Oy	181 TACTGCTATTTCCGACGCAAGTAGCGGATGCTATGACCACAACAACGCCACGACGATT 240	
Dd	181 TACTGCTATTTCCGCGCCGAATACGCGCATGATCATGACACAGGACCATGCCCCAGACCATC 240	
Oy	241 TGCGCGCGGATGACGCGCGCGCACGCCCTGCGCGCGCGGACGCTTGGCGCAACAATCACTAC 300	
Dd	241 TGCGCGCGGATGCATGCGCGCGCACGCCCTGCGCGCGCGGATGCGCGCAACAATCACTAT 300	
Oy	301 ACCGACTGGCGCGCGCAACAATTTCTATCCGCGCGCTGCGCGCAGCTGACACAGGCGCGCAAG 360	
Dd	301 ACCGACTGGCGCGCGCAACAATTTCTATCAACAGGCGCGCTGCGCGCAATCAATCCCAGCGCG 360	
Oy	361 CGGATCGGATCGAGTTGACACGTCGAATCTTGACTTCGCGCGCGACGCTTGAAGAACCC 420	
Dd	361 CGGATCGGATCGAGTTGATCACTGACGTCGAATCTTGACTTCGCGCGCGACGCTTGAAGAACCC 420	
Oy	421 CTACCGGGGCTGCACTTCTGTCGACATCAACCCAGCCCTGCAATTTGATGATGGGCAACATCAAG 480	
Dd	421 CTGCGCGGGCTGCAATTTCTGATATTCGATCAACCGCTGATATGATGGGCAACCGATCAAG 480	
Oy	481 TCAGCTGGAAG 540	
Dd	481 TCAGCTGGAAG 540	
Oy	541 GCCTCGCGCGCTGCAATCAAGGCGCGCGCTGCGCGGACGATGAAATGGCATGTGCACCACT 600	

Query Match	81.0%;	Score 981.6;	DB 6;	Length 1215;	
Best Local Similarity	88.1%;	Pred. No. 3.9e-116;			
Matches 1068;	Conservative 0;	Mismatches 144;	Indels 0;	Gaps 0;	
Qy	1	ATATCTGACGACATCTTTTGACCGTGTGATGAAATGGCCACACGCGCAGAAAGATTATTTGGCCG	60		
Db	1	ATATCTGACGACATCTTTTGACCGTGTGATGAAATGGCCACAAATGTGTAGAAAGAAATATTTCCCCC	60		
Qy	61	TTTTTCGGATCCGACAGATGACCCCGCCGCAAAAGACGTTTCGCGCTGAGATGGCCAGAAC	120		
Db	61	TTTTTCGGATCCGACAGATGACCGCCGCCGACAGATGACGTGGCGGCTGAGATGGCCAGAAAC	120		
Qy	121	AATGTGATGCGGCGCGCTGTTACCTTTATCACTGATCACTACTATTTCCGGCTGGCTG	180		
Db	121	GAGGTGACGCTCGCTGTTTCACTCTCTATTCATTTGATCAACTACTACTCTGTGATGGCTG	180		
Qy	181	TATCGGATTTTGGAGGCAAGTATGCGGCAATGATATGACACCAACAAAGCCACAGAGATT	240		
Db	181	TATGCTATTTTGGCCGCAATATGCGCAATGATATGACACAGAACATGCAAGACATATC	240		
Qy	241	TCGGCCGGCATTCGACGCGCGCCAGCCCTGTGGCGCCGACGTTTCGGCGACAAATCACTTAC	300		
Db	241	TCGGCCGGCATTCGATGGCGGTGAGCCCTGTGGCGCCGTAAGCTTGGCGCAACATCACTTAT	300		
Qy	301	ACCGACTTGGGCGCCGCGCAATTTCTATTCGCGCGCTGTCCGCAAGCTTACACAGGGCGCCAG	360		
Db	301	ACGAGCTTGGCGCGCGCAACATTTTACAGGCGCGTCCGTAATCACTCACCCCGCGCTAGG	360		
Qy	361	CGGATCGGATCGAGTTTCCACACGTCATCTTCGACTTCCGCGCGCAGCTTCGAGAAAGCC	420		
Db	361	CGGATCGGATTCGAGTTTCCATCAGCTGAGACCTTTGACTTCCGCGCGACGCTCGAAGAGGGG	420		
Qy	421	CTACCGGGCGTGCATCTTCGTGCACATCAGCCAGCCCTGCATGTGATGTCGACCATCAAG	480		
Db	421	CTGCGCGGCGTGCAGTTTCGTGCATATCGGTTCAACCGTCGATGTGATGTCGACGCTCAAG	480		
Qy	481	TGCGTGGAAAGAGAGAGCTGATTCGCGCAAGAGCGCCCGCTGTGTATGACGTTCGGCGCGGAG	540		
Db	481	TGCGTGGAAAGAGAGAGCTGATTCGCGAGGGGTGCCGTATCTTGCACGTTCCGCGGTGCC	540		
Qy	541	GCCTGCGCGCTGCATCAAGGCGCGGCTGCCGAGCATGAAATGGCATTCGCAACACAC	600		



Db 841 AAGACGTGCGGTGACCGCGCGGCTCGAATCATCAAGCCGGGTGCGGCTGCAAG 900  
Qy 901 GACATCGGCATCGAGCTCAACGAGATGTATCCGCCAGTGGACCTGTGTAAGTACCGCTCC 960  
Db 901 GATATCGGCATCGAATCTCAACGAGATGTATCCGCCAGTGGATGTGTGTAAGTACCGCTCC 960  
Qy 961 TTGGCTATGGGCACTCGCTTGGGGGTGCTGTGCACTATCACTGGTCCGGAGCGGGGCTG 1020  
Db 961 TTGGCTATGGGCACTCGCTTGGGGGTGCTGTGCACTATCACTGGTCCGGAGCGGGGCTG 1020  
Qy 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCCCGGACATGTGTCTCCATGAGACCG 1080  
Db 1021 GAACTGCGGAGGACATCGATATACGTGTGTCAGCCCGGACATGTGTCTCCATGAGACCG 1080  
Qy 1081 ATGTGTATGCTGCGCGGAGGAGATCCCGGTGCTCCGGCTATTCGGAGCACAGATCTCTG 1140  
Db 1081 ATGTGTATGCTGCGCGGAGGAGATCCCGGTGCTCCGGCTATTCGGAGCACAGATCTCTG 1140  
Qy 1141 ATCGTGGGAGGAGACGCGTGGCAGAACATCAACGGGCTTCCGCTTGGATCCGGGAAACAACC 1200  
Db 1141 ATCGTGGGAGGAGACGCGTGGCAGAACATCAACGGGATTCCTTGGGCTGAGCACAAAC 1200  
Qy 1201 ATCATCGGCAAC 1212  
Db 1201 ATCATCGGCAAC 1212

RESULT 12  
AX721943 1212 bp DNA linear PAT 07-MAY-2003  
LOCUS AX721943  
DEFINITION Sequence 11 from Patent EP1298213.  
ACCESSION AX721943  
VERSION AX721943.1 GI:30422519  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Shao, Z., Schmuck, R., Kratzsch, P., Kenklies, J. and Weisner, H.  
TITLE Variants of an erwinia-type CreA protein  
JOURNAL Patent: EP 1298213-A 11 02-APR-2003  
Roche Diagnostics GmbH (DE) ; F. HOFFMANN-LA ROCHE AG (CH)  
location/Qualifiers  
FEATURES  
source 1..1212  
1..1212  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant CT2m9"  
1..>1212  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAD90218.1"  
/db\_xref="GI:30422520"  
/translation="MTDDMLHMKHNGEKEYSPSDAMETRRQSDNRVEMENVDVA  
AFTSTHCHTNYSGFLICYTGRKRYKAVTDDHATTISAGIDGQPRKRSFGDITLTD  
WRNDNFYQAVRQLTPGARRIIGLEFDHVDLDFRTLEBALPVEFVILGQPSMMRITVK  
SLEBQCLIREGARI CDVGGA CVAAVAGVBEHVAIATNNAVRIASPFVEVLD  
TWTWFSGINTDGAHPVTNRI VOSGDIISLNTFPMI FGYYTALERTLFCDHVDSL  
DTWEKXVAHRRGLEL IKGARCKDIAIENEMYREMDLTKRSFGYGSFVLSHY  
GREAGVEAREDDITVIVQPGMVYSMEPMVLPBAPGAGGYREHDIIVGEDAENITG  
PFGPEPHNITRN"

ORIGIN  
Query Match 80.6%; Score 976.8; DB 6; Length 1212;  
Best Local Similarity 87.9%; Pred. No. 1.6e-115;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACATGTTGACGATGAAATGGCACACCGCGAGAAATATTATTCGCCG 60  
Db 1 ATGACTGACGACATGTTGACGATGAAATGGCACACATGTGTGAGAAATATTATTCGCCG 60  
Qy 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGACGTTTCGGGCTGATGGCCAAAGAC 120

Db 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGACGTTTCGGGCTGATGGCCAAAGAC 120  
Qy 121 AATGTGATGCGGCGGCTGTCACTCTTATCACTGATCAACTACTATTCGCGGTGCTG 180  
Db 121 GAGGTGACGCTGCGCTTCACTCTTATCACTGATCAACTACTATTCGCGGTGCTG 180  
Qy 181 TACTGCTATTTGCGAGCGAAGTAGGCGATGTATGACCAACAACGCCACAGCAATT 240  
Db 181 TACTGCTATTTGCGAGCGAAGTAGGCGATGTATGACCAACAACGCCACAGCAATT 240  
Qy 241 TTGGGCGGAGATCGAGCGGCGCGCGCTGGCGCGGAGCTTGGGCAACAATCACTAC 300  
Db 241 TTGGGCGGAGATCGAGCGGCGCGCGCTGGCGCGGAGCTTGGGCAACAATCACTAC 300  
Qy 301 ACCGATGCGCGCGCGCAATTTCTATTCGCGCGGCGCGCGAGCTGACACAGCGCGC 360  
Db 301 ACCGATGCGCGCGCGCAATTTCTATTCGCGCGGCGCGCGAGCTGACACAGCGCGC 360  
Qy 361 CGCATGCGATCGAGTTTCAACGTCATCTTGAATTCGCGCGCGAGCTGAGAAAGCC 420  
Db 361 CGCATGCGATCGAGTTTCAACGTCATCTTGAATTCGCGCGCGAGCTGAGAAAGCC 420  
Qy 421 CTACGCGGCGGTGCACTTGTGCAATCAACGCGCGCTGATGTGATGGGCAACATCAAG 480  
Db 421 CTACGCGGCGGTGCACTTGTGCAATCAACGCGCGCTGATGTGATGGGCAACATCAAG 480  
Qy 481 TCGCTCGAAGAGCAGAGCTGATTCGCGAAGCGCGCGGTGTGTACGTGCGCGCGCG 540  
Db 481 TCGCTCGAAGAGCAGAGCTGATTCGCGAAGCGCGCGGTGTGTACGTGCGCGCGCG 540  
Qy 541 GCGTGGCGGCTGCGCATCAAGCGCGCGGTGCGCGAGCATGAAGTGGCATCGCACACC 600  
Db 541 GCGTGGCGGCTGCGCATCAAGCGCGCGGTGCGCGAGCATGAAGTGGCATCGCACACC 600  
Qy 601 AATGCGATATCCGCGAGATTCGCGAATTCGCGTTCCTTGTGAGTGTGAGTGTGAC 660  
Db 601 AATGCGATATCCGCGAGATTCGCGAATTCGCGTTCCTTGTGAGTGTGAGTGTGAC 660  
Qy 661 ACTCGTTCAGTTCGCGCATCAACCGCGCGCGCGCGCATCAATTCGCGTCAACCGCATC 720  
Db 661 ACTCGTTCAGTTCGCGCATCAACCGCGCGCGCGCGCATCAATTCGCGTCAACCGCATC 720  
Qy 721 GTGCAATCCGCGGACATCTTGTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780  
Db 721 GTGCAATCCGCGGACATCTTGTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780  
Qy 781 GGGCTGGAGCGACGCTGTTCGCGACCATGTGCAATGAGCGACGCTGACATCTGGGAG 840  
Db 781 GGGCTGGAGCGACGCTGTTCGCGACCATGTGCAATGAGCGACGCTGACATCTGGGAG 840  
Qy 841 AAGAACTGCGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
Db 841 AAGAACTGCGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
Qy 901 GACATGCGCATGACGCTCAACGAGATGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960  
Db 901 GATATGCGCATGACGCTCAACGAGATGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960  
Qy 961 TTGGCTATGGGCACTCGCTTGGGGGTGCTGTGCACTATCACTGGTCCGGAGCGGGGCTG 1020  
Db 961 TTGGCTATGGGCACTCGCTTGGGGGTGCTGTGCACTATCACTGGTCCGGAGCGGGGCTG 1020  
Qy 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCCCGGACATGTGTCTCCATGAGACCG 1080  
Db 1021 GAACTGCGGAGGACATCGATATACGTGTGTCAGCCCGGACATGTGTCTCCATGAGACCG 1080  
Qy 1081 ATGTGTATGCTGCGCGGAGGAGATCCCGGTGCTCCGGCTATTCGGAGCACAGATCTCTG 1140  
Db 1081 ATGTGTATGCTGCGCGGAGGAGATCCCGGTGCTCCGGCTATTCGGAGCACAGATCTCTG 1140  
Qy 1141 ATCGTGGGAGGAGACGCGTGGCAGAACATCAACGGGCTTCCGCTTGGATCCGGGAAACAACC 1200

Db 1141 ATCGTGGGGGAAGCGGGCGAGAGAACTACCGGATTCCCTTCCGGCTGAGACACAC 1200  
Qy 1201 ATCATCGGCAAC 1212  
Db 1201 ATCATCGGCAAC 1212

RESULT 13  
AX721945 1212 bp DNA linear PAT 07-MAY-2003  
LOCUS Sequence 13 from Patent EPI298213.  
ACCESSION AX721945  
VERSION AX721945.1 GI:30422521  
KEYWORDS  
SOURCE Synthetic construct  
ORGANISM Synthetic construct  
REFERENCE 1  
AUTHORS Shao Z., Schmuck R., Kratzsch P., Kenklies J. and Weisser H.  
TITLE Variants of an erwinia-type creatinase  
JOURNAL Patent: EP 1298213-A 13 02-APR-2003;  
Roche Diagnostics GmbH (DE) ; F. HOFFMANN-LA ROCHE AG (CH)  
FEATURES  
source  
1. .1212  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant CT2m10"  
1. .>1212  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAD90219.1"  
/db\_xref="GI:30422522"  
/translation="MTDDMLVKKMHNGEKEYSPFDAMETRRSDYRRMAENDVDA  
ALFTSYHCINYSGLFVGRKXGMYIDODHATTISAGIDGGPMRSPGDNITYD  
WRBNFYQAVRQLTPGARIGIEFDHDLDEPRRLTEALIGVEVDIGOSMMARVTK  
SLEBOKIIRGARICDYGGAACVAAYAGPEHEVAATTNMAREIAKSPPELMD  
TWTFQSGINTDGMNPNVTRIVOSGDIISLANTPMIFGYTALERTLFCDHVDNLSL  
DTWKNVAVHRRGLEIKPGARKDIAELEINEMREMDLKYRSFGYSHSY  
GREAGVELREDIDVLQPGMVSMPEVMPLPEGAPGAGYREHDIIVGSDGABNITG  
PFLGPEHNITRN"

ORIGIN  
Query Match 80.6%; Score 976.8; DB 6; Length 1212;  
Best Local Similarity 87.9%; Pred. No. 1.6e-115;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACGATGTGACGTGATGAAGGACAAACGCGCGAAGATTATTCGCCG 60  
Db 1 ATGACTGACGACGATGTGACGTGATGAAGGACAAACGCGCGAAGATTATTCGCCG 60  
Qy 61 TTTTCGATGCGGAGATGACCGCGCGCAAAACGACGTTCCGCGCTGATGCGCAAGAC 120  
Db 61 TTTTCGATGCGGAGATGACCGCGCGCAAGATGACGCGCGCTGATGCGCGAAGAC 120  
Qy 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGACATCAACTACTGAGTTCCTG 180  
Db 121 GACGTGACGCTGCTGTTCACTCTTATCACTGACATCAACTACTGAGTTCCTG 180  
Qy 181 TACGTGATTTTGGACCGAAGTAGGCGATGCGATGCGACCAACGCGACGACGAT 240  
Db 181 TACGTGATTTTGGACCGAAGTAGGCGATGCGATGCGACCAACGCGACGACGAT 240  
Qy 241 TCGGCGCGGACATGACGCGCGCGACCGCTGCGCGCGACGCTTCCGCGACACATCCTAC 300  
Db 241 TCGGCGCGGACATGACGCGCGCGACCGCTGCGCGCGACGCTTCCGCGACACATCCTAT 300  
Qy 301 ACCGACTGCGCGCGCGACATTTCTATGCGCGCGCGCGACGCTGACGACGCGCGCGCAAG 360  
Db 301 ACCGACTGCGCGCGCGACACCTTCTACGCGCGCGCGCAACTACCGCGCGCGCGCGAG 360  
Qy 361 CGCATCGGACATGAGTTGACACGAGTCAATCTGACTTCCGCGCGCGCGCGCGAGAGCC 420

Db 361 CGCATCGGATTCGAGTTTCGATCAGGTGACCTTGACTTCCGCGCGCGCGCGAGAGCC 420  
Qy 421 CTACCGGCGCTGACTTCTGTCGACATCAGCCAGCCCTCGATGTGATGCGCACATCAAG 480  
Db 421 CTGCGCGCGCTGAGTTCTGTCGATTCGTCACCGCTCGATGTGATGCGCACATCAAG 480  
Qy 481 TCGCTGGAAGAGAGAAAGCTGATCCCGGAAGGCGCGCGCTGTGTATACCTGCGCGCGCG 540  
Db 481 TCGCTGGAAGAGAGAAAGCTGATCCCGGAAGGCGCGCGCTGTGTATACCTGCGCGCGCG 540  
Qy 541 GCCTGCGCGCTGCTCATCAAGCGCGCGCTGCGCGCGAGCATGAAGTGGCATGCGCACACC 600  
Db 541 GCCTGCGCTGCGCGCTCATCAAGCGCGCGCTGCGCGCGAGCATGAAGTGGCATGCGCACACC 600  
Qy 601 AATGCGATGATCGCGAGATTCGCAATGCTTCCCTTCTGTCGAGTGTATGACACCTGG 660  
Db 601 AATGCGATGATTCGCGAGATTCGCAATGCTTCCCTTCTGTCGAGTGTATGACACCTGG 660  
Qy 661 ACCTGCTTCCAGTTCGCGCATCAACACCGAGCGCGCGCGACAAATCCGCTCAACCGCATC 720  
Db 661 ACCTGCTTCCAGTTCGCGCATCAACACCGAGCGCGCGCGACAAATCCGCTCAACCGCATC 720  
Qy 721 GTGCAATCGGCGACATCTCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780  
Db 721 GTGCAATCGGCGATCTCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780  
Qy 781 GCGCTGAGCGACGCTGTTCTGTCGACATGTCGATGACGCGCGCGCATCATCTGGAG 840  
Db 781 GCGCTGAGCGAAGCGCTGTTCTGTCGACATGTCGATGACGCGCGCGCATCATCTGGAG 840  
Qy 841 AAGAACTGCGCGTGCATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900  
Db 841 AAGAACTGCGCGTGCATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900  
Qy 901 GACATGCGCATGACGCTCAACGAGATGTACCGCGAGTGGAGACCTGCTGAATGACCGCTCC 960  
Db 901 GATATGCGCATGACATCAACGAGATGTACCGCGAGTGGAGATGCTGTAAGTACCGCTCC 960  
Qy 961 TTGCGCTATGCGCATCTCTTCCGCGCTGCTGTGCACTACGCTGCGAGCGCGCGCTG 1020  
Db 961 TTGCGCTATGCGCATCTCTTCCGCGCTGCTGTGCACTACGCTGCGAGCGCGCGCGCTG 1020  
Qy 1021 GAGCTGCGGAGACATTCGACCGGAGCTGAAGCGCGCGAGTGTGCTTCATGAGCGCG 1080  
Db 1021 GAACTGCGGAGACATTCGATTCGCTGTCAGCGCGCGAGTGTGCTTCATGAGCGCG 1080  
Qy 1081 ATGGTATGCTGCGCGAGGACATGCGCGTGCAGCGCGCGCTATGCGAGACGACATCTTG 1140  
Db 1081 ATGGTATGCTGCGAGAGGCGCTCCCGCGCGCGCGCGCTATGCGAGACGACATCTTG 1140  
Qy 1141 ATGCTGCGGAGAGACGCTGCGAGAAACATACCGCTTCCGCTGCTGCGTGAACACACC 1200  
Db 1141 ATGCTGCGGAGAGACGCGCGCAGAGAAACATTAACGAGATTCCCTCCGCGCTGAGCAAC 1200  
Qy 1201 ATCATCGGCAAC 1212  
Db 1201 ATCATCGGCAAC 1212

RESULT 14  
AX721947 1215 bp DNA linear PAT 07-MAY-2003  
LOCUS Sequence 15 from Patent EPI298213.  
ACCESSION AX721947  
VERSION AX721947.1 GI:30422523  
KEYWORDS  
SOURCE Synthetic construct  
ORGANISM Synthetic construct  
REFERENCE 1  
AUTHORS Shao Z., Schmuck R., Kratzsch P., Kenklies J. and Weisser H.  
TITLE Variants of an erwinia-type creatinase

JOURNAL Patent: EP 1298213-A 15 02-APR-2003;  
Roche Diagnostics GmbH (DE) ; F.HOFMANN-LA ROCHE AG (CH)  
FEATURES  
Location/Qualifiers

1..1215  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant CT2m28"  
1..1215  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD90220.1"  
/db\_xref="GI:30422524"  
/translation="MTDDMLHVMKHNKEKEYSPSDAEMTRROSDVRBMAENDVDA  
ALFTSHICINTYSGFLYCYFGRKYGMYIDQHAATTISAGIDGQPMRBSFDNITYTD  
WRDNFYQAVRQLPFGARRIGIEFDHVDLDFRTELEALPGVEFVDIGQSPMWRITVK  
SLEBOKLIREGARI CDVGAACVAAGVAPHEVAIATYNAVRBIAASFPEVLAD  
TWTWFOGINTDGAHPVTRNRIVOSGDI SLNTPMIFGYTALERTLFCDHVDASL  
DTWKKVAVHRRGIELIKPGARKDIALENYREMDLKYRSFGYHSGFVLSHY  
GREAGVIREDDIVTLOPGMYVSNRPVMLPEGAPGAGTREDILIVGEDGAEINTG  
PFPEPHNITLRN"

## ORIGIN

Query Match 80.6%; Score 976.8; DB 6; Length 1215;  
Best Local Similarity 87.9%; Pred. No. 1.6e-115;  
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1 ATGATGACGACATGTTGACGTGATGAAATGACAAACGCGAGAAAGATTATTCGCG 60  
DB 1 ATGATGACGACATGTTGACGTGATGAAATGACAAATGTCATGTAAGAAATATTCGCG 60  
QY 61 TTTTCGATGCGGAGATGACCGCGCGCCAAAACAGCTTCGCGCTGATGCGCAAGAC 120  
DB 61 TTTTCGATGCGGAGATGACCGCGCGCCAAAACAGCTTCGCGCTGATGCGCAAGAC 120  
QY 121 AATGTCATGCGGCGCTGTTTCACTTATCACTGATCACTACTATTCGCGCTGCTG 180  
DB 121 GACGTCAACGCTGCGCTGTTTCACTTATCACTGATCACTACTATTCGCGCTGCTG 180  
QY 181 TACTGCTATTTGCGACGCAATGCGGATGCTGATCGACCAACCAACGCGACGAGATT 240  
DB 181 TACTGCTATTTGCGACGCAATGCGGATGCTGATCGACCAACCAACGCGACGAGATT 240  
QY 241 TCGGCGCGGATCGACCGCGCGCCAGCTTCGCGCGCGAGCTTCGCGGACATCACTAC 300  
DB 241 TCGGCGCGGATCGACCGCGCGCCAGCTTCGCGCGCGAGCTTCGCGGACATCACTAC 300  
QY 301 ACCGACTGCGCGCGGCAATTTCTATCGCGCGCTGCGCGCGAGCTGACCAACGCGCGCAG 360  
DB 301 ACCGACTGCGCGCGGCAATTTCTATCGCGCGCTGCGCGCGAGCTGACCAACGCGCGCAG 360  
QY 361 CGCATCGGATCGAGTTGACCAACGATCACTTCGCGCGCGAGCTTCGAGGAAGCC 420  
DB 361 CGCATCGGATCGAGTTGACCAACGATCACTTCGCGCGCGAGCTTCGAGGAAGCC 420  
QY 421 CTACCGGCGCTGCACTTCGTCGATCAACGAGCTTCGATGATGAGCGCACTCAAG 480  
DB 421 CTGCGCGCGCTGCACTTCGTCGATCAACGAGCTTCGATGATGAGCGCACTCAAG 480  
QY 481 TCGCTGAGAGGAGAGCTGATTCGGAAGGCGCGCGCTGCTGAGCGTTCGCGCGCGG 540  
DB 481 TCGCTGAGAGGAGAGCTGATTCGGAAGGCGCGCGCTGCTGAGCGTTCGCGCGCGG 540  
QY 541 GCCTGCGCGCTGCACTCAAGCGCGCGCTGCGCGAGATGAGTGGGATTCGCGCGCAGC 600  
DB 541 GCCTGCGCGCTGCACTCAAGCGCGCGCTGCGCGAGATGAGTGGGATTCGCGCGCAGC 600  
QY 601 AATGCGATGATCGCGGAGATGCGCAAAATCGTTCCCTTCGAGGCTGATGAGCACTTG 660  
DB 601 AATGCGATGATCGCGGAGATGCGCAAAATCGTTCCCTTCGAGGCTGATGAGCACTTG 660  
QY 661 ACCTGCTTCAAGTTCGCGGATCAACCGGCGCGGCAATTCGCGCTGCAACCAACCGCATC 720

DB 661 ACCTGCTTCAAGTTCGCGGATCAACCGGCGCGGCAATTCGCGCTGCAACCAACCGCATC 720  
QY 721 GTGCAATCGCGGACATCTTTTGCTTCAACACTTCCCGATGATCTTCGCTTACTACAC 780  
DB 721 GTGCAATCGCGGACATCTTTTGCTTCAACACTTCCCGATGATCTTCGCTTACTACAC 780  
QY 781 GCGTGGAGGCAACGCTGTTCTGAGACCATGATGATGAGCGCAAGCTTCGATCTGGAG 840  
DB 781 GCGTGGAGGCAACGCTGTTCTGAGACCATGATGATGAGCGCAAGCTTCGATCTGGAG 840  
QY 841 AAGAACGTGCGCGTGCATCGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900  
DB 841 AAGAACGTGCGCGTGCATCGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900  
QY 901 GACATCGCATCGAGCTCAACGAGATGATCGCGAGTGGAGCTTCGTAAGTACCGCTCC 960  
DB 901 GATATCGCTTCGAACCAACGAGATGATCGCGAGTGGAGTGTGCTGAAGTACCGCTCC 960  
QY 961 TTCGCGTATGCGCACTCTTCGCGCGCTGCTGCACTGACGCTCGCGAGCGCGCGGTG 1020  
DB 961 TTCGCGTATGCGCACTCTTCGCGCGCTGCTGCACTGACGCTCGCGAGCGCGCGGTG 1020  
QY 1021 GACCTGCGCGGACATCGACACCGAGCTGAGCCCGCGCATGCTGCTTCATGAGCCG 1080  
DB 1021 GACCTGCGCGGACATCGACACCGAGCTGAGCCCGCGCATGCTGCTTCATGAGCCG 1080  
QY 1081 ATGATGATGCTGCGCGGAGATGCGCGGCTGCGCGGCTATGCGACGACGATCTCTG 1140  
DB 1081 ATGATGATGCTGCGCGGAGATGCGCGGCTGCGCGGCTATGCGACGACGATCTCTG 1140  
QY 1141 ATGCTGGGAGAGACGCGCGAGAACATTCGAGATTCCTTCGCGCTTGAACGACAC 1200  
DB 1141 ATGCTGGGAGAGACGCGCGAGAACATTCGAGATTCCTTCGCGCTTGAACGACAC 1200  
QY 1201 ATCATCGGCAAC 1212  
DB 1201 ATCATCGGCAAC 1212  
RESULT 15  
AX721953 1212 bp DNA linear PAT 07-MAY-2003  
LOCUS AX721953  
DEFINITION Sequence 21 from Patent EP1298213.  
ACCESSION AX721953  
VERSION AX721953.1 GI:30422529  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Shao,Z., Schmuck,R., Kratzsch,P., Kenklies,J. and Weisner,H.  
AUTHORS  
TITLE Variants of an erwinia-type crealtnase  
JOURNAL Patent: EP 1298213-A 21 02-APR-2003;  
Roche Diagnostics GmbH (DE) ; F.HOFMANN-LA ROCHE AG (CH)  
FEATURES  
Location/Qualifiers  
1..1212  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant CT6d7"  
1..1212  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD90223.1"  
/translation="MTDDMLHVMKHNKEKEYSPSDAEMTRROSDVRBMAENDVDA  
ALFTSHICINTYSGFLYCYFGRKYGMYIDQHAATTISAGIDGQPMRBSFDNITYTD  
WRDNFYQAVRQLPFGARRIGIEFDHVDLDFRTELEALPGVEFVDIGQSPMWRITVK  
SLEBOKLIREGARI CDVGAACVAAGVAPHEVAIATYNAVRBIAASFPEVLAD  
TWTWFOGINTDGAHPVTRNRIVOSGDI SLNTPMIFGYTALERTLFCDHVDASL  
DTWKKVAVHRRGIELIKPGARKDIALENYREMDLKYRSFGYHSGFVLSHY  
GREAGVIREDDIVTLOPGMYVSNRPVMLPEGAPGAGTREDILIVGEDGAEINTG  
PFPEPHNITLRN"

GRGAGVELBEDIDTVLQPGMNV/SMEPMVLPBGPAGAGSYREHDLIYEDGAEINITG  
FPLGPRHNIRN"

Query Match 80.5%; Score 975.2; DB 6; Length 1212;  
Best Local Similarity 87.8%; Pred. No. 2,6e-115;  
Matches 1064; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAATGAGCAACGCGAGAAAGATTATTCGCG 60  
DB 1 ATGACTGACGACATGTTGCACTGATGAATGAGCAACGCGAGAAAGATTATTCGCG 60  
QY 61 TTTTGAGATCGGAGATGACCGCGCGCAAAAGAGTTCCGCGCTGATGTCGCAAGAC 120  
DB 61 TTTTGAGATCGGAGATGACCGCGCGCAAAAGAGTTCCGCGCTGATGTCGCAAGAC 120  
QY 121 AATGTGATGCGCGCGCTGTTCACTTCTTATCACTGATCACTACTATTCCGAGTGG 180  
DB 121 GAGTGTGACGCTGCGCTGTTCACTTCTTATCACTGATCACTACTACTGATGCTG 180  
QY 181 TACTGTATTGTTGAGACGACGATGCGCATGTCATGACCAACACCGCACGAGATT 240  
DB 181 TACTGTATTGTTGAGACGACGATGCGCATGTCATGACCAACACCGCACGAGATT 240  
QY 241 TGGGCGGCGATGACGCGCGCGACGCGCTGCGCGCGAGCTTGGCGGCAACATCACTAC 300  
DB 241 TGGGCGGCGATGACGCGCGCGACGCGCTGCGCGCGAGCTTGGCGGCAACATCACTAT 300  
QY 301 ACCGACTGCGCGCGCGACAAATTTCTATCGCGCGTGCAGCTGACCAAGCGCGCAAG 360  
DB 301 ACCGACTGCGCGCGCGACAAATTTCTATCGCGCGTGCAGCTGACCAAGCGCGCAAG 360  
QY 361 CGCATGCGATGAGTTGTCATGATGACGCTTGAACCTTCCGCGCACGCTCGAAGAGCG 420  
DB 361 CGCATGCGATGAGTTGTCATGATGACGCTTGAACCTTCCGCGCACGCTCGAAGAGCG 420  
QY 421 CTACCGGCGCTGACCTTCTGACATCAACGCGCTGATGTCGATGTCGACATCAAG 480  
DB 421 CTACCGGCGCTGACCTTCTGACATCAACGCGCTGATGTCGATGTCGACATCAAG 480  
QY 481 TCGCTCGAAGACGAGCTGATCCGCGAGCGCGCGCTGTCGACGTCGCGCGCGCG 540  
DB 481 TCGCTCGAAGACGAGCTGATCCGCGAGCGCGCGCTGTCGACGTCGCGCGCGCG 540  
QY 541 GCTGCGCGCGCTGCTCAAGCGCGCGCTGCGCGAGCAAGAGTGCATCGCCACAC 600  
DB 541 GCTGCGCGCGCTGCTCAAGCGCGCGCTGCGCGAGCAAGAGTGCATCGCCACAC 600  
QY 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTGTGAGAGTGAACCTGAG 660  
DB 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTGTGAGAGTGAACCTGAG 660  
QY 661 ACCTGCTTCAGTCCGCGCATCAACCGCGCGCGCAATCCGCTCAACACCGCATC 720  
DB 661 ACCTGCTTCAGTCCGCGCATCAACCGCGCGCGCAATCCGCTCAACACCGCATC 720  
QY 721 GTGCAATCGCGCGCACTCTTGTGCTCAACCTTCCGATGATCTTGGCTACTAC 780  
DB 721 GTGCAATCGCGCGCACTCTTGTGCTCAACCTTCCGATGATCTTGGCTACTAC 780  
QY 781 GCGCTGAGCGCGACGCTTGTGCGACATGTCGATGACGCGAGCTGACATCTGGAG 840  
DB 781 GCGCTGAGCGCGACGCTTGTGCGACATGTCGATGACGCGAGCTGACATCTGGAG 840  
QY 841 AAGAACTGCGCGCTGATCGCGCGCGCTGAGCTGATCAACCGCGCGCGCTGCAAG 900  
DB 841 AAGAACTGCGCGCTGATCGCGCGCGCTGAGCTGATCAACCGCGCGCGCTGCAAG 900  
QY 901 GACATCGCGCTGAGCTCAAGAGATGTCGCGAGTGGGACCTGCTGAAGTACCGCTCC 960  
DB 901 GATATCGCGCTGAGCTCAAGAGATGTCGCGAGTGGGATCTGCTGAAGTACCGCTCC 960  
QY 961 TTGCGCTATGCGCACTCTTCCGCGGCTGTGTCACACTACTACGCTCGGAGCGCGG 1020

DB 961 TTTGCTTACGCGCACTCTTCCGCTGCTCTCCACTACTACGCGCGCGAGCGGCGCTG 1020  
QY 1021 GAGTGTGCGAGACATGACACCGAGCTGAAGCCCGCATGCTGCTCCATGAGCGG 1080  
DB 1021 GAGTGTGCGAGACATGACATGCTGCTGAGCGCGCGCATGCTGCTCCATGAGCGG 1080  
QY 1081 ATGCTGATGCTGCGGAGGCGATGCGCGCTGCGCGCTATCGGAGCAAGACATCTCTG 1140  
DB 1081 ATGCTGATGCTGCGGAGGCGCTGCGCGCTGCGCGCTATCGGAGCAAGACATCTCTG 1140  
QY 1141 ATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 1141 ATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1201 ATCATCGCGCAC 1212  
DB 1201 ATCATCGCGCAC 1212

Search completed: November 19, 2004, 18:11:19  
Job time : 5279 secs